

L5 ANSWER 1 OF 1 GENBANK.RTM. COPYRIGHT 2001

LOCUS (LOC): BF509252 GenBank (R)
GenBank ACC. NO. (GBN): **BF509252**
CAS REGISTRY NO. (RN): 307875-04-3
SEQUENCE LENGTH (SQL): 791
MOLECULE TYPE (CI): mRNA; linear
DIVISION CODE (CI): Expressed sequence tag
DATE (DATE): 6 Dec 2000
DEFINITION (DEF): UI-H-BI4-aow-c-07-0-UI.s1 NCI_CGAP_Sub8 Homo sapiens
cDNA clone IMAGE:3086220 3', mRNA sequence.
KEYWORDS (ST): EST
SOURCE: human.
ORGANISM (ORGN): Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Euteleostomi; Mammalia; Eutheria; Primates;

Catarrhini;

Hominidae; Homo

NUCLEIC ACID COUNT (NA): 270 a 151 c 144 g 226 t

COMMENT:

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: M13 Forward

POLYA=Yes.

REFERENCE: 1 (bases 1 to 791)

AUTHOR (AU): NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE (TI): National Cancer Institute, Cancer Genome Anatomy
Project (CGAP), Tumor Gene Index

JOURNAL (SO): Unpublished (1997)

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..791	/organism="Homo sapiens" /db-xref="taxon:9606" /clone="IMAGE:3086220" /clone-lib="NCI-CGAP-Sub8" /lab-host="DH10B (Life Technologies)" /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site-1: Not I; Site-2: Eco RI; NCI-CGAP-Sub8 is a subtracted library derived from NCI-CGAP-Sub5. The NCI-CGAP-Sub8 library had 2.5 million recombinants. A single-stranded DNA preparation of NCI-CGAP-Sub5 was used as a tracer in a subtractive hybridization with a driver comprising: a pool of clones from NCI-CGAP-Sub5 (IMAGE

clone Ids 2732833-2737415,
 3068040-3069191; 25% of the driver
 population), a pool of clones from
 NCI-CGAP-Sub4 (IMAGE clone Ids
 2723592-2729326; 25% of the driver
 population), NCI-CGAP-Sub6 (pool
 AIF-AJU, IMAGE Ids
 2728969-2733190; 25% of the driver
 population), and NCI-CGAP-Sub7
 (IMAGE Ids 3069192-3072238,
 3081864-3084550 ; 25% of the
 driver population). Subtraction
 was performed as previously
 described [Bonaldo, Lennon &
 Soares (1996): Normalization and
 Subtraction: Two Approaches To
 Facilitate Gene Discovery. Genome
 Research 6, 791-806.
 TAG-LIB=NCI-CGAP-Lu5
 TAG-TISSUE=lung TAG-SEQ=CAAC"

SEQUENCE (SEQ):

```

1 tttttttttt tttgttacca tgaaatcatc acattagaac aatatcagct atgtgcatta
61 gctatcaacg atgaatgcta tcaagtaaga caagtgtttg cccagaaact tcacaaaggc
121 ctttcccgtt tacggcttcc acttgagtat atggcaatct gtgccctttg tgcaaaagat
181 cctgtaaagg agagaagagc tcatgctagg caatgtttgg tgaaaaatat aaatgtaagg
241 cgggagtatc tgaagcagca tgcagctggt agtgaaaaat tattgtctct tctaccagag
301 tatgttggtc catatacaat tcaccttttg gcacatgacc cagattatgt caaagtacag
361 gatattgaac aacttaaaga tgttaaagaa tgtcttttgt ttgttctgga aatattaatg
421 gctaaaaatg aaaataacag tcacgctttt atcagaaaaga tggtagaaaa tattaaacia
481 acaaaagatg cccaaggacc agatgatgca aaaatgaatg aaaaactgta cactgtgtgt
541 gatgttgcca tgaatatcat catgtcaaag agtactacat acagtttgga atctcctaaa
601 gacccggtac taccagctcg tttcttctact caacctgaca agatttcagt aacacaaaaa
661 attatctgcc tcctgaaatg aaatcatttt tcactcctgg aaaacctaaa acaacaatgt
721 tctaggagct gttaacaagc actttcatca gcaggcaagc atctcagaca gatcatcaga
781 atctcgtgcc g

```

=>

L3 ANSWER 1 OF 1 GENBANK.RTM. COPYRIGHT 2001

LOCUS (LOC): HSU95825 GenBank (R)
GenBank ACC. NO. (GBN): U95825
CAS REGISTRY NO. (RN): 225494-81-5
SEQUENCE LENGTH (SQL): 5253
MOLECULE TYPE (CI): mRNA; linear
DIVISION CODE (CI): Primates
DATE (DATE): 5 Apr 1999
DEFINITION (DEF): Human androgen-induced prostate proliferative shutoff
associated protein (AS3) mRNA, complete cds.
SOURCE:
ORGANISM (ORGN): human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Euteleostomi; Mammalia; Eutheria; Primates;

Catarrhini;

Hominidae; Homo

NUCLEIC ACID COUNT (NA): 1755 a 944 c 1074 g 1480 t

COMMENT:

On Apr 5, 1999 this sequence version replaced gi:4539617.

REFERENCE: 1 (bases 1 to 5253)

AUTHOR (AU): Geck, P.; Szelei, J.; Jimenez, J.; Soto, A.M.;
Sonnenschein, C.

TITLE (TI): Androgen-induced proliferative shutoff in prostate
cancer cells

JOURNAL (SO): Proc. Annu. Meet. Am. Assoc. Cancer Res., 37, 223-223
(1996)

REFERENCE: 2 (bases 1 to 5253)

AUTHOR (AU): Geck, P.; Szelei, J.; Jimenez, J.; Sonnenschein, C.;
Soto, A.M.

TITLE (TI): Early gene expression during androgen-induced
inhibition of proliferation of prostate cancer cells:

a

new suppressor candidate on chromosome 13, in the
BRCA2-Rb1 locus

JOURNAL (SO): J. Steroid Biochem. Mol. Biol., 68 (1-2), 41-50 (1999)
OTHER SOURCE (OS): CA 131:3561

REFERENCE: 3 (bases 1 to 5253)

AUTHOR (AU): Geck, P.; Szelei, J.; Jimenez, J.; Sonnenschein, C.;
Soto, A.M.

TITLE (TI): Direct Submission

JOURNAL (SO): Submitted (28-MAR-1997) Anatomy and Cell Biology,

Tufts

University Medical School, 136 Harrison Avenue,

Boston,

MA 02111, USA

REFERENCE: 4 (bases 1 to 5253)

AUTHOR (AU): Geck, P.; Szelei, J.; Jimenez, J.; Sonnenschein, C.;
Soto, A.M.

TITLE (TI): Direct Submission

JOURNAL (SO): Submitted (05-APR-1999) Anatomy and Cell Biology,

Tufts

University Medical School, 136 Harrison Avenue,

Boston,

MA 02111, USA

FEATURES (FEAT):

Feature Key Location Qualifier

=====+=====+=====

source	1..5253	/organism="Homo sapiens" /db-xref="taxon:9606" /chromosome="13" /map="13q12-q13" /cell-line="LNCaP" /cell-type="cancer cells" /tissue-type="prostate" /note="exon 1 maps to P1 Artificial Chromosome PAC26H23; exons 2-24 map to cosmid 267p19; exons 25-34 map to P1 Artificial Chromosome PAC 49J10"
gene	1..5253	/gene="AS3"
exon	1..46	/gene="AS3"
		/number=1
exon	47..173	/gene="AS3"
		/number=2
CDS	66..4241	/gene="AS3"
		/codon-start=1
		/product="androgen-induced prostate proliferative shutoff associated protein"
		/protein-id="AAD22134.2"
		/db-xref="GI:4559410"
		/translation="MAHSKTRTNDGKITYPPGVK EISDKISKEEMVRRLLKMVVKTFMD MDQDSEEEKELYLNLLALHLASDFFLKHPDKDVRL LVACCLADIFRIYAPEAPYTSPOK LKDIFFMITRQLKGLDTSKSPQFNRYFYLLNIA WVKSYNICFELEDSNEIFTQLYRT LFSVINNGHNQKVMHMMVDLMSSIIICGDTVSQE LLDTVLVNLVPAHKNLNKQAYDLA KALLKRTAQAIPEYITNFFNQVLMGKTSISDLS EHVSDLILELYNIDSHLLSVLPQ LEFKLKSNDNEERLQVVKLLAKMFGAKDSELASQ NKPLWQCYLGRFNDIHVPIRLECV KFASHCLMNHPDLAKDLTEYLKVRSHDPEEAIKH DVIVSIVTAAKKDILLVNDHLLNF VGERTLDRWRVRKEAMMGLAQIYKKYALQSAAG KDAKQIAWIKDKLLHIYYQNSID DRLLVERIFAQYMPHNLETTERMKCLYYLYATL DLNAVKALNEMWKCQNLLRHQVKD LLDLIKQPKTDASVKAFISKVMVITRSLPDGKA QDFMKKFTQVLEDDKIRKQLEVL VSPTCSCQAEQCVREITKKLGNPKQPTNPFLEM IKFLLERIAPVHIDTESISALIKQ VNKSIDGTADDEDEGVPTDQAIRAGLELLKVLSE THPISFSAETFESELLACLKMDDE KVAEALQIFKNTGSKIEEDFPHIRSALLPVLHH KSKKGPPRQAKYAIHCHAIHFSSK EAQFAQIFEPLHKS LDPSNLEHLITPLVTIGHIA LLAPDQFAAPLKS LVATFIVKDLL MNDRLPGKTTKLWVPDEEVSPETMVKIQAIKMM VRWLLGMKNHNSKSGTSTLRLTT ILHSDGDLTEQKGKISKPDMSRLRLAAGSAIVKLA QEPYHEIITLEQYQLCALAINDE CYQVRQVFAQKLHKGLSRLRPLEYMAICALCAK DPVKERRAHARQCLVKNINVRREY LKQHAHVSEKLLSLLPEYVVPYTIHLLAHDPDYV KVQDIEQLKDVKECLWFVLEILMA

exon 174..377
 exon 378..464
 exon 465..562
 exon 563..689
 exon 690..770
 exon 771..911
 exon 912..1027
 exon 1028..1122
 exon 1123..1268
 exon 1269..1420
 exon 1421..1534
 exon 1535..1616
 exon 1617..1665
 exon 1666..1805
 exon 1806..1921
 exon 1922..2027
 exon 2028..2188
 exon 2189..2312
 exon 2313..2471
 exon 2472..2540
 exon 2541..2677
 exon 2678..2801
 exon 2802..3006
 exon 3007..3121

KNENNSHAFIRKMVENIKQTKDAQGPDDAKMNEK
 LYTVCDVAMNIIMSKSTTYSLESP
 KDPVLPARFFTQPDKNFSNTKNYLPPGMKSFFTP
 GKPKTTNVLGAVNKPLSSAGKQSQ
 TKSSRMETVGNASSSSNPSSPGRIKGRLDSSSEM
 HSENYDTMSSPLPGKRS DKRDDS
 DLVRSELEKPRGRKKTPTVTEREEKLGMDLTKLV
 QEQRPKGSQSRKRGHTASESEDEQ
 QWPEEKRLKEDILENEDEQNSPPKKGKRGRPPKP
 LGGGTPKEEPTMKTSKKGSKKKSG
 PPAPEEEEEERQSGNTEQKSKSKQHRVSRRAQQ
 RAESSESAIESTQSTPQKGRGRP
 SKTPSPSQPKKNV"
 /gene="AS3"
 /number=3
 /gene="AS3"
 /number=4
 /gene="AS3"
 /number=5
 /gene="AS3"
 /number=6
 /gene="AS3"
 /number=7
 /gene="AS3"
 /number=8
 /gene="AS3"
 /number=9
 /gene="AS3"
 /number=10
 /gene="AS3"
 /number=11
 /gene="AS3"
 /number=12
 /gene="AS3"
 /number=13
 /gene="AS3"
 /number=14
 /gene="AS3"
 /number=15
 /gene="AS3"
 /number=16
 /gene="AS3"
 /number=17
 /gene="AS3"
 /number=18
 /gene="AS3"
 /number=19
 /gene="AS3"
 /number=20
 /gene="AS3"
 /number=21
 /gene="AS3"
 /number=22
 /gene="AS3"
 /number=23
 /gene="AS3"
 /number=24
 /gene="AS3"
 /number=25
 /gene="AS3"

exon	3122..3254	/number=26 /gene="AS3"
exon	3255..3374	/number=27 /gene="AS3"
exon	3375..3437	/number=28 /gene="AS3"
exon	3438..3583	/number=29 /gene="AS3"
exon	3584..3689	/number=30 /gene="AS3"
exon	3690..4129	/number=31 /gene="AS3"
exon	4130..4354	/number=32 /gene="AS3"
exon	4355..4810	/number=33 /gene="AS3"
		/number=34

SEQUENCE (SEQ) :

1	ccggagagcc	ccggagtgag	cggagtagcg	agtcggcaac	ccggaggggt	agaaatattt
61	ctgtcatggc	tcattcaaag	actaggacca	atgatggaaa	aattacatat	ccgcctgggg
121	tcaaggaaat	atcagataaa	atatctaaag	aggagatggt	gagacgatta	aagatggttg
181	tgaaaacttt	tatggatatg	gaccaggact	ctgaagaaga	aaaggagctt	tatttaaacc
241	tagctttaca	tcttgcttca	gatttttttc	tcaagcatcc	tgataaagat	gttcgcttac
301	tggtagcctg	ctgccttgct	gatattttca	ggattttatgc	tcctgaagct	ccttacacat
361	cccctgataa	actaaaggat	atattttatgt	ttataacaag	acagttgaag	gggctagagg
421	atacaaagag	cccacagttc	aataggtatt	tttatttgct	tgagaacatt	gcttgggtca
481	agtcataata	catatgcttt	gagttagaag	atagcaatga	aattttcacc	cagctataca
541	gaaccttatt	ttcagttata	aacaatggcc	acaatcagaa	agtccatatg	cacatggtag
601	accttatgag	ctctattatt	tgtgaaggta	atacagtgtc	tcaggagctt	ttggatacgg
661	tttttagtaaa	tctgggtacct	gctcataaga	atttaaaca	gcaagcatat	gatttggcaa
721	aggctttact	gaaggaggaca	gctcaagcta	ttgagccata	tattaccaat	ttttttaatc
781	aggttctgat	gcttgggaaa	acatctatca	gcgatttgct	agagcatgtc	tctgacttaa
841	ttttggagct	ctacaatatt	gatagtcatt	tgctgctctc	tgttttaccc	cagcttgaat
901	ttaaattaaa	gagcaatgat	aatgaggagc	gcctacaagt	tgttaaacta	ctggcgaaaa
961	tgtttggggc	aaaggattca	gaattggctt	ctcaaaaaca	gccactttgg	cagtgtctact
1021	tgggcagggt	taatgatatc	catgtaccaa	tccgcctgga	atgtgtgaaa	tttgctagcc
1081	attgtctcat	gaaccatcct	gatttagcaa	aagacttaac	agagtatctt	aaagtgaggt
1141	cacatgaccc	tgaggaagct	attagacatg	atgttattgt	gtcaatagtt	acagctgcta
1201	aaaaggatat	tcttctggtc	aatgatcact	tacttaattt	tgtgggagag	agaacattag
1261	acaaacgatg	gagagtacgc	aaagaagcca	tgatgggact	tgcccaaatt	tataagaaat
1321	atgctttaca	gtcagcagct	ggaaaagatg	ctgcaaaaaca	gatagcatgg	atcaaagaca
1381	aattgctaca	tatatattat	caaaaatagta	ttgatgatcg	actacttggt	gaacggatct
1441	ttgctcaata	catggttcct	cacaatttag	aaactacaga	acggatgaaa	tgcttatatt
1501	acttgtatgc	cacactggat	ttaaattgctg	tgaaagcatt	gaatgaaatg	tggaaatgtc
1561	aaaatctgct	ccgacatcaa	gtaaaggatt	tgcttgactt	gattaagcaa	cccaaaaacag
1621	atgccagtg	caaggccata	ttttcaaaa	tgatggttat	tacaagaagt	ttacctgatc
1681	ctggtaaggc	tcaggatttc	atgaagaaat	tcacacaggt	gttagaagat	gatgagaaaa
1741	taagaaagca	gttagaagta	cttgttagtc	caacatgctc	ctgcaagcag	gctgaaggtt
1801	gtgtgctgta	aataactaag	aagttagggca	accccaaaaca	gcctacaaat	cctttcctgg
1861	aaatgatcaa	gtttctcttg	gagaggatag	cacctgtgca	catagatacc	gaatctatca
1921	gtgctcttat	taaacaagtg	aacaaatcaa	tagatggaac	agcagatgat	gaagatgagg
1981	gtgttccaac	tgatcaagcc	atcagagcag	gtccttgaact	gcttaaggta	ctctcattta
2041	cacatcccat	ctcatttcat	tctgctgaaa	catttgaatc	attactggcc	tgtctgaaaa
2101	tgtagatgta	aaaagtagca	gaagctgcac	tacaaatttt	caaaaacaca	ggaagcaaaa
2161	ttgaagagga	ttttccacac	atcagatcag	ccttgcttcc	tgttttacat	cacaaatcta
2221	aaaaaggacc	cccccgtaaa	gccaaatatg	ccattcattg	tatccatgcy	atattttcta
2281	gtaaagaggc	ccagtttgca	cagatatttg	agcctctgca	taagagccta	gatccaagca
2341	acctggaaca	tctcataaca	ccattgggtta	ctattgggtca	tattgctctc	cttgcacctg
2401	atcaatttgc	tgctcctttg	aaatccttgg	tagctacttt	cattgtgaaa	gatcttctca

2461	tgaatgatcg	gcttccaggg	aaaaagacaa	ctaaactttg	ggttccagat	gaagaagtat
2521	ctcctgagac	aatgggtcaaa	attcaggcta	ttaaaatgat	ggttcgatgg	ctacttggaa
2581	tgaaaaataa	tcacagtaaa	tcaggaaactt	ctaccttaag	attgctaaca	acaatattgc
2641	atagtgatgg	agacttgaca	gaacagggga	aaattagtaa	accagatatg	tcacgtctga
2701	gacttgctgc	tgggagtgct	attgtgaagc	tggcacaaga	accctgttac	catgaaatca
2761	tcacattaga	acaatatcag	ctatgtgcat	tagctatcaa	cgatgaatgc	tatcaagtaa
2821	gacaagtgtt	tgcccagaaa	cttcacaaa	gcctttcccg	tttacggctt	ccacttgagt
2881	atatggcaat	ctgtgccctt	tgtgcaaaa	atcctgtaaa	ggagagaaga	gctcatgcta
2941	ggcaatgttt	ggtgaaaaat	ataaatgtaa	ggcgggagta	tctgaagcag	catgcagctg
3001	ttagtgaaaa	attattgtct	cttctaccag	agtatgttgt	tccatataca	attcaccttt
3061	tggcacatga	cccagattat	gtcaaagtac	aggatattga	acaacttaaa	gatgttaaag
3121	aatgtctttg	gtttgttctg	gaaatattaa	tggctaaaaa	tgaaaataac	agtcacgctt
3181	ttatcagaaa	gatggtagaa	aatattaagc	aaacaaaaga	tgcccaagga	ccagatgatg
3241	caaaaatgaa	tgaaaaactg	tacactgtgt	gtgatgttgc	catgaatatc	atcatgtcaa
3301	agagtactac	atacagtttg	gaatctccta	aagacccggg	actaccagct	cgtttcttca
3361	ctcaacctga	caagaatttc	agtaacacca	aaaattatct	gcctcctgga	atgaaatcat
3421	ttttcactcc	tggaaaacct	aaaacaacca	atgttctagg	agctgttaac	aagccacttt
3481	catcagcagg	caagcaatct	cagaccaaat	catcacgaat	ggaaactgta	ggcaatgcaa
3541	gcagcagctc	aaatccaagc	tctcctggaa	gaataaaggg	gaggcttgat	agttctgaaa
3601	tggatcacag	tgaaaatgaa	gattacacaa	tgtcttcacc	tttgccgggg	aaaagaagtg
3661	acaagagaga	cgactctgat	cttgtaaggt	ctgaattgga	gaagcctaga	ggcagaaaaa
3721	aaacgcccgt	cacagaacgg	gaggagaaat	taggtatgga	tgacttgact	aagttgggtac
3781	aggaacagag	acctaaggc	agtcagcgaa	gtcggaaaag	aggccatacg	gcttcagaat
3841	ctgatgaaca	gcagtggcct	gaggaaaaga	ggctcaaaga	agatatatta	gaaaatgaag
3901	atgaacagaa	tagtccgcca	aaaaagggta	aaagaggccg	accaccaaaa	cctcttggtg
3961	gaggtacacc	aaaagaagag	ccaacaatga	aaacttctaa	aaaaggaagc	aaaaaaaaat
4021	ctggacctcc	agcaccagag	gaggaggaag	aagaagaaa	acaaagtgga	aatacggaac
4081	agaagtccaa	aagcaaacag	caccgagtg	caaggagagc	acagcagaga	gcagaatctt
4141	ctgaatctag	tgcaattgaa	tccacacagt	ccacaccaca	gaaaggacga	ggaagaccat
4201	caaaaacgcc	atcaccatca	caaccaaaaa	aaaatgtgta	agttgtaaat	attacatttc
4261	aaaccaattt	caaattattt	tgcaaaagtt	cctaaatttg	taaacatata	tattgctgta
4321	tttaaattcc	atatattttag	ccccattaca	ctaggtacgg	cggcgaagtg	ctaaaaggga
4381	acggcgatga	acaaatgtaa	ttaataactt	tctctgtgaa	agctttggaa	aaatcttttt
4441	tttttttttt	tttttttggt	caagcttgag	gctgaataaa	gcctttgatg	cacaaaatgg
4501	gactgctgaa	gagtggacag	ttggacctta	ctttggtgac	cccatacatt	tgtggtcaca
4561	tgcttttagcc	atacacatgg	taacattgac	tatggagtct	tgtgaaagtg	taatgtgcga
4621	tggctatgta	gacataaaga	agaaacttgt	aaatatcttt	tttctttttt	ttaatgtttc
4681	tgattttctga	agtgtctgta	tagcttttat	ctgcggcttt	aaactgacag	taccgcagctg
4741	tttattggat	ctattgattt	gaaaagaatt	tgttaggata	gatcttaagc	agtaatctgt
4801	cagtgtttgt	atttgtattt	tctgcaattt	tactgtgaaa	aaaaatttgt	tttcaacaat
4861	tgggtgtcatt	ttcttgatgt	cactattttgt	tggagagtta	aatgggtctct	tccctttgtg
4921	tatcttacct	agtgtttact	cctgggcacc	cttaatcttc	agaggtgcta	aattgtctgc
4981	cattacacca	gaaggatgcc	tctgatagga	ggacaacat	gcaaattgtg	aaatagtcct
5041	gaagttcttg	gattacttta	cacctcagta	ttgatttgtc	ccagaatttt	ctggcctttc
5101	atggcaatga	aaatttttaag	aagaaaagatt	taaagtattt	taatttttaa	gagtgtgtta
5161	taaaataatg	tactgaattc	tttatcccat	tttatcatcc	tttcagtttt	tattaatcta
5221	ctgtatcaat	aaaattctgt	aatttgaatg	agt		

=>

L3 ANSWER 1 OF 1 GENBANK.RTM. COPYRIGHT 2001

LOCUS (LOC): AC068224 GenBank (R)
GenBank ACC. NO. (GBN): **AC068224**
CAS REGISTRY NO. (RN): 263690-85-3
SEQUENCE LENGTH (SQL): 54398
MOLECULE TYPE (CI): DNA; linear
DIVISION CODE (CI): High Throughput Genome Sequencing
DATE (DATE): 30 Apr 2000
DEFINITION (DEF): Homo sapiens chromosome 3 clone RP11-660H19 map 3,
LOW-PASS SEQUENCE SAMPLING.
KEYWORDS (ST): HTG; HTGS_PHASE0
SOURCE: human.
ORGANISM (ORGN): Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Euteleostomi; Mammalia; Eutheria; Primates;

Catarrhini;

Hominidae; Homo

NUCLEIC ACID COUNT (NA): 13631 a 10245 c 8645 g 14936 t 6941 others
COMMENT:

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L10161

Center clone name: 660_H_19

* NOTE: This record contains 68 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1 685: contig of 685 bp in length
* 686 785: gap of 100 bp
* 786 1513: contig of 728 bp in length
* 1514 1613: gap of 100 bp
* 1614 2354: contig of 741 bp in length
* 2355 2454: gap of 100 bp
* 2455 3184: contig of 730 bp in length
* 3185 3284: gap of 100 bp
* 3285 3959: contig of 675 bp in length
* 3960 4059: gap of 100 bp
* 4060 4746: contig of 687 bp in length
* 4747 4846: gap of 100 bp
* 4847 5564: contig of 718 bp in length
* 5565 5664: gap of 100 bp
* 5665 6360: contig of 696 bp in length
* 6361 6460: gap of 100 bp


```

*      6461      7132: contig of 672 bp in length
*      7133 7232: gap of      100 bp
*      7233      7937: contig of 705 bp in length
*      7938 8037: gap of      100 bp
*      8038      8749: contig of 712 bp in length
*      8750 8849: gap of      100 bp
*      8850      9522: contig of 673 bp in length
*      9523 9622: gap of      100 bp
*      9623     10356: contig of 734 bp in length
*     10357 10456: gap of      100 bp
*     10457     11181: contig of 725 bp in length
*     11182 11281: gap of      100 bp
*     11282     11975: contig of 694 bp in length
*     11976 12075: gap of      100 bp
*     12076     12785: contig of 710 bp in length
*     12786 12885: gap of      100 bp
*     12886     13557: contig of 672 bp in length
*     13558 13657: gap of      100 bp
*     13658     14347: contig of 690 bp in length
*     14348 14447: gap of      100 bp
*     14448     15132: contig of 685 bp in length
*     15133 15232: gap of      100 bp
*     15233     15912: contig of 680 bp in length
*     15913 16012: gap of      100 bp
*     16013     16731: contig of 719 bp in length
*     16732 16831: gap of      100 bp
*     16832     17506: contig of 675 bp in length
*     17507 17606: gap of      100 bp
*     17607     18339: contig of 733 bp in length
*     18340 18439: gap of      100 bp
*     18440     19109: contig of 670 bp in length
*     19110 19209: gap of      100 bp
*     19210     19900: contig of 691 bp in length
*     19901 20000: gap of      100 bp
*     20001     20693: contig of 693 bp in length
*     20694 20793: gap of      100 bp
*     20794     21464: contig of 671 bp in length
*     21465 21564: gap of      100 bp
*     21565     22245: contig of 681 bp in length
*     22246 22345: gap of      100 bp
*     22346     23049: contig of 704 bp in length
*     23050 23149: gap of      100 bp
*     23150     23857: contig of 708 bp in length
*     23858 23957: gap of      100 bp
*     23958     24685: contig of 728 bp in length
*     24686 24785: gap of      100 bp
*     24786     25499: contig of 714 bp in length
*     25500 25599: gap of      100 bp
*     25600     26302: contig of 703 bp in length
*     26303 26402: gap of      100 bp
*     26403     27108: contig of 706 bp in length
*     27109 27208: gap of      100 bp
*     27209     27910: contig of 702 bp in length
*     27911 28010: gap of      100 bp
*     28011     28720: contig of 710 bp in length
*     28721 28820: gap of      100 bp
*     28821     29484: contig of 664 bp in length
*     29485 29584: gap of      100 bp
*     29585     30328: contig of 744 bp in length
*     30329 30428: gap of      100 bp

```

```

* 30429 31111: contig of 683 bp in length
* 31112 31211: gap of 100 bp
* 31212 31931: contig of 720 bp in length
* 31932 32031: gap of 100 bp
* 32032 32710: contig of 679 bp in length
* 32711 32810: gap of 100 bp
* 32811 33502: contig of 692 bp in length
* 33503 33602: gap of 100 bp
* 33603 34301: contig of 699 bp in length
* 34302 34401: gap of 100 bp
* 34402 35100: contig of 699 bp in length
* 35101 35200: gap of 100 bp
* 35201 35914: contig of 714 bp in length
* 35915 36014: gap of 100 bp
* 36015 36709: contig of 695 bp in length
* 36710 36809: gap of 100 bp
* 36810 37518: contig of 709 bp in length
* 37519 37618: gap of 100 bp
* 37619 38330: contig of 712 bp in length
* 38331 38430: gap of 100 bp
* 38431 39109: contig of 679 bp in length
* 39110 39209: gap of 100 bp
* 39210 39912: contig of 703 bp in length
* 39913 40012: gap of 100 bp
* 40013 40712: contig of 700 bp in length
* 40713 40812: gap of 100 bp
* 40813 41514: contig of 702 bp in length
* 41515 41614: gap of 100 bp
* 41615 42294: contig of 680 bp in length
* 42295 42394: gap of 100 bp
* 42395 43094: contig of 700 bp in length
* 43095 43194: gap of 100 bp
* 43195 43915: contig of 721 bp in length
* 43916 44015: gap of 100 bp
* 44016 44740: contig of 725 bp in length
* 44741 44840: gap of 100 bp
* 44841 45533: contig of 693 bp in length
* 45534 45633: gap of 100 bp
* 45634 46330: contig of 697 bp in length
* 46331 46430: gap of 100 bp
* 46431 47137: contig of 707 bp in length
* 47138 47237: gap of 100 bp
* 47238 47927: contig of 690 bp in length
* 47928 48027: gap of 100 bp
* 48028 48709: contig of 682 bp in length
* 48710 48809: gap of 100 bp
* 48810 49509: contig of 700 bp in length
* 49510 49609: gap of 100 bp
* 49610 50331: contig of 722 bp in length
* 50332 50431: gap of 100 bp
* 50432 51135: contig of 704 bp in length
* 51136 51235: gap of 100 bp
* 51236 51981: contig of 746 bp in length
* 51982 52081: gap of 100 bp
* 52082 52783: contig of 702 bp in length
* 52784 52883: gap of 100 bp
* 52884 53580: contig of 697 bp in length
* 53581 53680: gap of 100 bp
* 53681 54398: contig of 718 bp in length.

```

REFERENCE: 1 (bases 1 to 54398)

AUTHOR (AU): Birren,B.; Linton,L.; Nusbaum,C.; Lander,E.
 TITLE (TI): Homo sapiens chromosome 3, clone RP11-660H19
 JOURNAL (SO): Unpublished
 REFERENCE: 2 (bases 1 to 54398)
 AUTHOR (AU): Birren,B.; Linton,L.; Nusbaum,C.; Lander,E.;
 Abraham,H.; Allen,N.; Anderson,S.; Baldwin,J.;
 Barna,N.; Bastien,V.; Beda,F.; Boguslavkiy,L.;
 Boukhgalter,B.; Brown,A.; Burkett,G.; Campopiano,A.;
 Castle,A.; Choepel,Y.; Colangelo,M.; Collins,S.;
 Collymore,A.; Cooke,P.; DeArellano,K.; Dewar,K.;
 Diaz,J.S.; Dodge,S.; Domino,M.; Doyle,M.; Ferreira,P.;
 FitzHugh,W.; Gage,D.; Galagan,J.; Gardyna,S.;
 Ginde,S.;
 Goyette,M.; Graham,L.; Grand-Pierre,N.; Grant,G.;
 Hagos,B.; Heaford,A.; Horton,L.; Howland,J.C.;
 Iliev,I.; Johnson,R.; Jones,C.; Kann,L.; Karatas,A.;
 Klein,J.; LaRocque,K.; Lamazares,R.; Landers,T.;
 Lehoczký,J.; Levine,R.; Lieu,C.; Liu,G.; Locke,K.;
 Macdonald,P.; Marquis,N.; McCarthy,M.; McEwan,P.;
 McGurk,A.; McKernan,K.; McPheeters,R.; Meldrim,J.;
 Meneus,L.; Mihova,T.; Miranda,C.; Mlenga,V.;
 Morrow,J.;
 Murphy,T.; Naylor,J.; Norman,C.H.; O'Connor,T.;
 O'Donnell,P.; O'Neil,D.; Olivar,T.M.; Oliver,J.;
 Peterson,K.; Pierre,N.; Pisani,C.; Pollara,V.;
 Raymond,C.; Riley,R.; Rogov,P.; Rothman,D.; Roy,A.;
 Santos,R.; Schauer,S.; Severy,P.; Spencer,B.;
 Stange-Thomann,N.; Stojanovic,N.; Subramanian,A.;
 Talamas,J.; Tesfaye,S.; Theodore,J.; Tirrell,A.;
 Travers,M.; Trigilio,J.; Vassiliev,H.; Viel,R.; Vo,A.;
 Wilson,B.; Wu,X.; Wyman,D.; Ye,W.J.; Young,G.;
 Zainoun,J.; Zimmer,A.; Zody,M.
 TITLE (TI): Direct Submission
 JOURNAL (SO): Submitted (30-APR-2000) Whitehead Institute/MIT Center
 for Genome Research, 320 Charles Street, Cambridge, MA
 02141, USA

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..54398	/organism="Homo sapiens" /db-xref="taxon:9606" /chromosome="3" /map="3" /clone="RP11-660H19" /clone-lib="RPCI-11 Human Male BAC"

SEQUENCE (SEQ):

```

1 ggacccatgg gcgatatgttc tannnntggt ngagttagn cttccccctt cttccttaaa
61 atatgccacc tcccaacagc aaaacaaggc ggctccactg ccaggatttc atcatccctc
121 catccccatt tggggacggg agactatggt cgtcgacaga attttcttaa tttaaaattg
181 cccccactca cgctgtaat ccagcactt tgggtattata taaatgattt ataaatgata
241 taattcagat gtcatcattt tccagacagg tttgagggcc actggtgaaa ggggtggtgat
301 ggatgtcagg ggtgtggttc ccgtgggcat tttgtgctgc aaaggaaaga aggagaggaa
361 gagttaggaa aaatctgaga ttatgaaaag cagcaacatt tgggggttgt ttgaagaagt
421 gaggaataat ttagggatca aaattttagg gacaggtgaa gctggcgggc gtaataagga
481 cttccaaaat gcctttacca gccacagcag tggccagtgc tcagaaaagg caaccaacgc
541 tgagcaaggc tgctgccaag cccaggggat gaaccgctgg cgaaggtgaa atggcctcga
601 acagccaggg ggacttgtct atgttccaag ctccctcaag ggctccaggg ggactcctca

```

661	gccggtcctc	atccccctgc	atcccnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
721	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
781	nnnnngcaac	tggttgcnaag	aatccccctac	tttgtgttnan	agaataccctt	ttcattatgtt
841	gggatagata	ctttctaaac	acatgggtgac	tatgggttaat	aatactgttta	ttatatataac
901	tgaaatttgc	taagagagca	gagatcaagt	gttgtcacca	cacacacaca	cgagtaacta
961	ggtgaggtgt	ggatatgtta	gcttgattgt	ggcaaccatt	tcacaatata	tatgtatatac
1021	aaaatatcac	attgtacact	tcaaattgtat	acacttttta	tttgtcaagt	atacctcaat
1081	aaagctggag	aaaaaaataa	taataaagag	taagaccctt	ggaaatgcct	tgaggtactt
1141	ggtgacagtc	aatcttctga	cctcatctat	atccacaacc	caggttatat	ttctgtttga
1201	gagaataagg	atgtattttt	gtttactgga	ttatagtgcc	agcattagtc	ttactctgt
1261	ctccttctct	gtctctctct	ctctctctct	ctctctctct	ctctgtatct	agctctgctt
1321	ccctctctct	ccatctccct	ccatccatct	agtaggaata	ttgtgctggc	cttgggttcc
1381	atgctctctc	tcccctatgg	ctccctcact	ctgcctcggc	ctgagcgagc	atgaatcact
1441	tctgctgggg	tgtgctctgc	tgtactgctt	ttgttctcag	agctgngatg	taaatacgac
1501	tacttcagtg	tttnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
1561	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnngtggttg
1621	gggggaacaa	actattgctg	acaccccaac	acttctgtna	gnagaaaacc	cccataaact
1681	ttaattgtag	gaggggcctc	taacattgga	caaactttta	agatgtattt	tcatgaaagg
1741	tacttcttta	atggaatttg	agataattag	agcatttatg	atgacattta	attttgtgaa
1801	aagtatagtg	agcattttta	tgtaattttt	catttagagt	tcatgaaaga	ttttctgtgg
1861	aattggagca	tcagggaaat	tatggcaaag	aatgggcagc	cttattctga	atatctactg
1921	catggtttta	aagctgtgat	tgtgttttat	tttagagcaa	tgataatgag	gagcgccctac
1981	aagttgttaa	actactggca	aaaaattgtt	gggcaaagga	ttcagaattg	gcttctcaaa
2041	acaagccact	ttggcagtg	tacttgggca	ggtatatgat	tttggtttcc	catttaaaag
2101	taatacatta	ttctactgac	cgttttttta	agaaacatta	tccacattta	aactgtgttc
2161	tctggaatca	taaaaatgaa	ggatttttta	tgttgaaagt	tttgaaagtt	gttaatgact
2221	ggctatatat	atattttact	tttttagtct	ttatcctcag	taatttttaa	aagataatac
2281	attttaaagc	tatctcttat	ggtattatga	ccatgaagtt	aatattatta	ctttaaggca
2341	gtgattaatt	ccaannnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
2401	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnngtctctg
2461	gnngnggtaa	catggttaaa	agctgtacct	cgttanagtt	aacactncta	cttagaggca
2521	ttattcattc	tttttttccc	tcctctcaac	cgtcaccaaa	ttctgttaat	actttctcca
2581	aatgtcttga	acatttatct	tctactctcc	taccttgcta	tattttcact	tggattgttg
2641	tatacattgc	ctccctgttg	taattgttcc	tgtgcctctg	accactgcc	tgctgtatca
2701	cattgctatc	aaaattatcc	ttctaaaaca	gattagatca	catcagtcct	ttacttgaaa
2761	tttctgttga	ttctctgttg	cttgtgggat	aaagttgaaa	caccttatag	ccctttatag
2821	tttctctctc	gctgagttt	caagctttct	ttctgctccc	tattatgaaa	tctgtgttct
2881	atttactttc	gattacttga	aatttgtgta	acagcttttt	ttctctgttg	tcaactctat
2941	ctggaaattg	cttcctgttt	ccagacaaaac	aaacaatgaa	caaaaccaag	gaacatggtt
3001	gacacctttt	gaatgcttta	aaatacagtt	ttgggggttac	cacatcactg	tgaccctctt
3061	tctttttttt	tccttcctta	cgctggaaaa	tgtgaggggt	ctgcagtcac	atcccacttc
3121	tgccgtttac	ctgtcgtatt	ttggcagatt	acttaatcat	ttgtgcccta	gtttccttat
3181	ttgcnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
3241	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnacataa	ctggtgcaaa
3301	aaccgacaac	tcgttagaga	aaactgctga	ttttgattga	gatcatggga	atttgcctag
3361	gaaaagtga	agttttcacc	ccttagatga	aaacaacact	gctccattga	cctctgtcct
3421	atgaagtgt	tgtgtggaag	ggagattgag	aaaaagcact	ctaaaactgaa	aaagaaaaaa
3481	gaatactctt	gttttccctg	aaggcaattt	aaaatgacca	ggcgactgc	tttttagatg
3541	aattatgcct	tttacatgtc	tacatacagc	atttcaatga	aaaaaaaaaa	accctagaat
3601	cattctaaat	gattgataaa	agataaaaata	tgatacatcc	aatggatgaa	taccatcatt
3661	acattataaa	tactaggtag	tacattaaaa	tcttcataaa	tattaagtaa	aaacatcaag
3721	atacctaact	ttaattacat	aattattgaa	actttctaaa	accacaagca	tatacaaaaa
3781	agtagggaac	atcttagagt	ggcaacataa	agtgaatttt	ctttcgattt	tctatgcttt
3841	caataatgct	gagttattgt	tacaccttta	ttttaaagct	acaaaaaaac	tttgagtaca
3901	cagaccaca	ctctctgtaa	tgttgggtat	taccttttgc	agatataaac	atttatcacn
3961	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
4021	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	atactatttn	gcaaaccctc
4081	cacacttgct	tnanagaatc	ccctatatga	gatgggtgtt	aagccttaat	ttttgtatct
4141	tatttaagtt	agggcagatt	tatattttact	agtttttttc	tgttttaaaa	gtatgtttgt
4201	tttctaacac	atctaactgg	tacttattag	ggccagatg	aatattgcag	cttcaaaatg

4261	aacttttagtc	acaacatctc	actcttatga	gcagtgaagc	aattattgta	gataaacttg
4321	ctactgttta	ctctttactt	aaaatagcaa	attttatatc	tgaagttggt	actctttgct
4381	gccatttgaa	taatatctgg	aataattgtg	tctttacagg	atgaatgcta	tcaagtaaga
4441	caagtgtttg	cccagaaact	tcacaaaggc	ctttcccggt	tacggcttcc	acttgagtat
4501	atggcaatct	gtgccctttg	tgcaaaagat	cctgtaaaagg	agagaagagc	tcatgctagg
4561	caatgttttg	tgaaaaatat	aaatgtaagg	cgggagtatc	tgaagcagca	tgcagctggt
4621	agtggtaagc	atataagaaa	atggaaagga	tactttttca	gcctgctagt	ttcagtttta
4681	tagaatatag	catgatatat	gatattactg	ttattttatta	tttgtgatta	acaattttta
4741	ttaccannnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
4801	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnngaa	tgagggtgctg
4861	agtctgacga	tcggctctna	nagaaaacac	tttcatcagc	aggcaagcaa	tctcagacca
4921	aatcatcacg	aatggaaact	gtaagcaatg	caagcagcag	ctcaaatcca	agctctcctg
4981	gaagaataaa	ggggaggtaa	gtgcaaaaga	aatgccacaa	tttacattta	aggattccag
5041	caaggctatg	agatctgaaa	taatataatc	caaaatgttt	catgtcagta	aatggaaagt
5101	aatgtattac	ttaatttttc	catttaacat	tggtagtac	aagttcatac	ttatttgctt
5161	gctcattttg	gatgggtgct	ctgtcttagt	ctttcattct	ttagcttgct	gggggtgatg
5221	tgggggctg	ggcgtctccg	cgcggcggtg	cctggccgtg	ccccgctccc	cgtgtcctgc
5281	ccgtcttcgt	ctttgcgcgt	tcccttccct	cccgtttttt	tgcccttggt	cccctctccc
5341	cttttcttcg	tttcccttct	cccccccccc	cgttccctcc	ctttcccccg	tttttctttc
5401	cgcgtccctt	ttccctcgcc	ctttgtgttt	ttgccctggt	tccttccctt	tgccctttcc
5461	ccttccccct	tccccccctt	ccctcccccc	cgttctcctt	ttccttggtg	ccccccccct
5521	ccctcgtggt	cgtcccccg	ttggcccccc	cccgtttttc	ccgcnnnnnn	nnnnnnnnnn
5581	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
5641	nnnnnnnnnn	nnnnnnnnnn	nnnaaacact	tcgttgctga	cctccaaccc	tgactctana
5701	gnatccccc	tacttcttca	cttacacagt	cttctttggc	catgtcatca	tctatgtaag
5761	tgactattac	atcagcaacg	ttaagggttt	tcaagactcg	ggtcctactt	caagtgctaa
5821	acagaaatth	tatactttta	tgttctgtga	atataatact	gagagttata	atatttccca
5881	gatctattct	ttcttgacat	gtagagccta	tgtaagtgtt	aaatatagta	gttcacagat
5941	ctattctttc	ttgacatgaa	gagcctatat	aagtgttaaa	taaactctgaa	tggcatttct
6001	tatacctaaa	agcatgggtt	atatttgact	cttggcctgt	ccttgacact	tcttgcaaa
6061	tacagataat	tctgtccttt	gatattccatt	atataatth	tattcttact	ttcttgacca
6121	ttcttattgc	tgtaatcatc	attagctctc	accaagactg	ttgggaatct	tttttattgt
6181	tggtaaatth	ccctaattac	agttatgatc	atgtcattcc	tctgttaaaa	tatttagggg
6241	gttgatgctt	gtaaaaacaa	ctgcacgtat	tttacaacct	ggtgtttgag	cccctcta
6301	atttttaacc	aatgtacgc	ttaaaaatct	tggtgagcac	actctccttt	acttaacctn
6361	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
6421	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	gactactttt	gcagacaccc
6481	accctgctn	ctnanagaaa	cccccttcag	tttttgtttt	ttttctgacc	ttttcttgag
6541	ttcactcgct	tcagtthtgt	ctagctgtac	ctaaagctca	tgctttgaa	tcttaatttc
6601	acatatatac	acatacttca	attctgta	tctatagttt	ttaaacatag	ctctttgttt
6661	tgctgcagaa	tttgcttgct	attaaatcct	tgaaaaatct	taattatgtg	aagtttgat
6721	ctgttaggct	tatcatcttg	atctcttgta	acttttgtct	attctcaaaa	ttttggtcaa
6781	atcttacctt	gtacatatat	tgcttttgat	tgagtgatag	actttttgta	tgaaaagttg
6841	ttgagataat	gtgaggcttc	agatgtggtt	atattcttcc	agagttttac	ttttgtttct
6901	atcaggcagt	taagctatcg	tcagatcact	tccatccaat	gaaaatggaa	gtttggcttc
6961	agttttggac	aagcttatta	tttttttctg	gtcaccttat	gtctagagta	tatttgtctg
7021	ggatccaccc	agaagtccag	gttgcttaac	aaggccttta	cttaatgggc	tcaggattct
7081	taatcttggt	cttctagctc	tgtgacagga	aagctttgct	caacttctta	tnnnnnnnn
7141	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
7201	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnacaacttc	ttgctgacac	tgccatactt
7261	gacgctgana	gnatccccc	ctaaacan	anaaccacat	taaaactaaa	agtcacaaag
7321	ggtacataaa	gcacagcaaa	ggaaacaatc	aacaaaatga	aaagcagcct	ataggatggg
7381	aaaaatattt	gcaaaccatg	tttccaacaa	ggaattaatg	tcttaaggta	cacgaggact
7441	catacaactc	agtagtaaca	aaagccaaat	cacccaataa	aaaatgtgca	aaggacctga
7501	gtagatattt	ttccaaagaa	gacataataa	taggaaagag	gtatatgtaa	agatgtcact
7561	aatcatcagt	gaaatagaaa	tcaaaaccac	aatgatatgt	cacttcatac	ctgtagatg
7621	gctattacaa	aaacattaaa	agttaagtgt	cagtgaggat	gtggaggaaa	gggatccctt
7681	atacatthtt	ggtaggaacg	taaattggtg	cagccattat	gaaaaacagc	gtggagatat
7741	cgcagaagct	taaaaataga	actaccctgt	gatccaacaa	tcccacgtct	gagtatatat
7801	ccaaaggaaa	tgaaattact	atcttgaaga	aatatctgac	ccttcatggt	catctcagca

7861	ttattcacag	aggccaagat	ttggaaacaa	cataaatatt	agtcagtggg	tgaatggata
7921	aagaaaatgt	agtgtgannn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
7981	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnngtt
8041	gttggnnccg	ataacttggt	gcagaaactn	acaacgtgtc	tanagaaaac	ctagaccctt
8101	ggccacaaca	tttttatcat	ccagtttaata	tacagccttg	ttttatgtat	gtttctgttt
8161	aaagactatt	tattatata	agttgtctca	actgaactca	cagccaacag	cactgtaact
8221	catacctgaa	caaagcttat	ctagcacaca	cattttctct	gtaaggaata	tcagagcctt
8281	cttgtgcttt	ggaacgccag	acagcagttc	atcgctatgc	ttgtgctttg	aaacaccaga
8341	cagtacttca	tcatagaaata	acaaaaacat	caaccacaaa	actgtgaaaa	atgtggcact
8401	gaatagacct	caaaaaggac	actgttttacg	tatgaaagct	gaaacaagaa	ggcacagtgt
8461	tgccttcttt	gacattagct	gggaatgtgc	acatctggct	acttacattt	tttgcgtctt
8521	tgtgtattta	ctttgcaa	gacccaaaaa	ttgcttcaag	tattaattnt	tggtgttaca
8581	aataaatttt	aacgagtatg	cagattccta	aatatagaat	ccacaaataa	ggaagactga
8641	ctgtatgtat	tcagttcctg	aaaggaaaac	aacagtatcg	aaatttatac	ttttataaag
8701	acatgacata	gaactaattt	caggttagaa	gcattcctta	ttactcagnn	nnnnnnnnnn
8761	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
8821	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnng	gacccttgtg	ctngaggctc	gactggtcta
8881	nagnaacact	tgtgtgggac	agctggccgc	ttcaaaacgc	tgacggaaca	gtagctgcaa
8941	tacaacggct	atcacgttgt	gtttgctctt	gcgggaagcc	cgaaagacgc	tgatgacaca
9001	tcaatctaca	tgttttatca	aaaggtcggc	gacaactcaa	tcgacagctg	gaaaaacgcg
9061	ggccgtgtct	ttaaagacag	cgataagttc	gacgccaacg	atccgatcct	gaaagatcag
9121	acgcaagaat	ggtccggttc	tgcaactctt	acatctgacg	gaaaaatccg	tttattctac
9181	ctgtactatt	ccggtaaaca	ttacggcaaa	caaagcctga	caacagcgca	ggtaaagtgt
9241	tcaaaatctg	atgacacact	caaaatcaac	ggagtggaa	atcacaaaac	gatttttgac
9301	ggagacggaa	aaacatatca	gaacgttcag	cagtttatcg	atgaaggcaa	ttatacatcc
9361	ggcgacaacc	atacgtgag	agaccctcac	tacgtttgaa	gacaaagcca	taaattccct
9421	tggatttcca	agcccacccg	ggaaccgaaa	aacggtccca	aggggaaaga	atcttttttt
9481	taacaaaggg	ttcttcgggg	ggggggccca	aacttttttc	cannnnnnnn	nnnnnnnnnn
9541	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
9601	nnnnnnnnnn	nnnnnnnnnn	nnggactaag	ctgtgctaga	aacccccagt	ccttactgtn
9661	agnagnaana	cccctaactt	ttgnaacaaa	attaccaana	gacatcatag	gtgtgatctg
9721	gctctttctt	atctcctttc	tatcctaccc	cttccttcag	cttttgcctc	attcaactcc
9781	agaaacatca	gccttcttga	aatccttgat	gcacaaagtt	cattcttgct	taaggacttt
9841	catacttatt	cttcctctg	cctggaatac	tcttctccca	gatattcata	aggcttgatg
9901	cctcacttaa	ttgaagtcca	tgctcagatg	ttaccttctc	aaagacgccg	cctttgacta
9961	ctctgtgtca	aatagcattc	ctgctgggga	gctgggcgga	atagcctgaa	tgaaatctcc
10021	tgagggttg	ctgtgctagt	ctgtcatata	atgcatgtct	taccattaag	acatttgaaa
10081	caaccagatt	gagtctaact	aaaggtgcaa	cccaggctca	actcttgaag	tcatcaactt
10141	ctagaatgcc	taaaagaaaa	aagggtcaat	ctgctcttgg	gagaaaaaat	acatcaactt
10201	caatccttta	gtacttttta	ttcaatgtcc	tcaaataaga	aaaaagttac	aagatccaca
10261	caaggaagga	aaaaatgtga	ccaggaaaag	tttgatggat	gttcaataga	aaaaagagtt
10321	aaaagaatca	gatcgacaaa	tgattgagaa	gttgannnn	nnnnnnnnnn	nnnnnnnnnn
10381	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
10441	nnnnnnnnnn	nnnnnnnacac	ttttgctgac	actgctaaac	ttctgttnana	gnatccaccc
10501	ctcaattagn	atatgtgcag	ctgtccacaa	aatcccagct	ttctcattat	taccctatag
10561	tctcacagca	cataagacat	catacaataa	atgtaaatac	ttgagtgaat	ttaatatctt
10621	tccaatagaa	atagaaattt	taatacctca	aatattacaa	tgtgggtgatt	atattaaaaa
10681	caatagatct	atgcacatta	ctcactttga	ccattgtctc	aggagatact	tcttcatctg
10741	gaacccaaag	tttagttgtc	tttttccctg	gaagctaaaa	ataaacgaga	aaaaaaaaaa
10801	aaacaaggac	tttaaaatgt	aaatatgaat	tcttttcaat	tagcacttaa	aaaatgaatg
10861	aaatctgtac	cttctgcagt	gacatttggt	atatatgact	aaagtacttc	atctaaccct
10921	atttacttta	aaaaaaaaaa	aagggttgaaa	ggcatacctt	cactagcgcc	tatgtacagc
10981	acttatagat	ccatcacgac	tggttctttc	taggtctgtg	gggggactaa	ctctaacttc
11041	ttacacatca	ctcccttatt	tagccctttt	tgagattcaa	ttatttaaatg	caagaactag
11101	gattttaaaa	aaacatggag	ccaaaagggtg	tagctcttaa	acaagaagaa	aggtataaaa
11161	aaaaaaaaatc	ttgagctgga	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
11221	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
11281	nacactgttg	ctgaacccca	gtttgtctct	anagnacccc	cccctacact	ttaaatcatc
11341	tctagattac	ttataatacc	taaatgctat	gtagtacta	tacagcattg	ctttttattt
11401	gtattatttg	tatatatttt	ctccaaatat	ttttgattcg	agttagttaa	atccccagat

11461	gcagaaactg	tagatctgga	aggctgatgg	tactcatctc	caaactcatc	agattgtata
11521	cattaaatat	gtacagggtt	ttatatctca	atcataactc	aaaaagtgg	tttttttta
11581	aaaaaaaaa	gaaagggtt	gcataaaatt	gacacagata	ttccatctga	attggagttg
11641	agttttgccc	aacattcctt	accacattgt	ccatgaaaaa	tcaactgttt	gagaatataa
11701	atgtaatat	tgcttccaaa	taccaagaaa	gataaggata	agaattttat	aagaaaagata
11761	aataatttga	tcattagttt	tttaaaggcc	tctaaatgcc	atgaggaagc	acaagcaa
11821	ccctgtaaga	ttttagggca	ggaagggatg	atttcaggat	gagatcatca	ggggagactc
11881	cagggtggtga	tgagcctttg	agctgggccc	tgaagaagag	gacggatttc	cataggcaga
11941	gacgaagctg	taatagaaag	cgtcagaggc	agagnnnnnn	nnnnnnnnnn	nnnnnnnnnn
12001	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
12061	nnnnnnnnnn	nnnnnaacta	acgtgctgaa	cccccgctcg	acnctnanag	natcccccta
12121	gcttcttttg	gctataacca	tgacttactc	atctttgtat	ttccagaact	tagaacaatg
12181	tacagcacac	agaagtacag	agtaaatctt	tcatgaactg	aattccaagt	tctaaaatat
12241	ctaagtaatt	attaaattat	gttaaaattc	atattgtttt	attaaatttc	tccaaatgta
12301	ctttctgttc	aaatacataa	cagcagcctc	ccagatacaa	tactgcattt	aaagagcctc
12361	aagaaataaa	gtacatgtac	taaagcgcaa	agcatagtcc	tgtacgtatc	agaggactgc
12421	agataaaata	aaagactcag	agcaagtttt	gtaaagaagg	ttgtgtctga	gcactgactt
12481	gcagaaaggg	ttaggttgga	tgtgtggaaa	tgggagagta	aagagaaaat	taaaggcagc
12541	aaagtacaag	gcacaaaaaa	gaaaaaaagc	tgttcagcta	catgaaatga	ggttaagtaa
12601	aggagagtgt	gctcaacaag	atttttaagc	gtacatttgg	gttaaaatat	tagagggggt
12661	caaacaccag	gttgtaaaat	acgtgcagtt	gtttttacaa	gcacaaacc	cctacatat
12721	ctaacagatg	aatgacatga	tcatcactgc	tgtagggag	atttaccacc	aactataaaa
12781	gatcannnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
12841	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnntacaa	catacttgca
12901	gaccccgctg	tcttggtcnc	tagnagaatc	cccttntatc	atgttgatac	aaacctaata
12961	gtattctaat	aagtagccca	acanagttca	tgtaattaac	caagctgttg	ctttctgcta
13021	aaatctaagg	atgtctccag	acttcccgtg	gtaaaaagtt	taccttacia	aatgcaagtt
13081	ttaaacagcc	agctataaatt	gtttgaatgg	tgagattgcc	atttcagagt	ttcaatgtat
13141	tgatctgaca	atctcatttc	aaactgacct	aagaaaactg	aactctgaaa	ggccatgttg
13201	cctgttgaa	acagtagtac	atcctttacat	tataggacaa	caaattgcaa	cgatctctaa
13261	aaaatgcttc	aaaactttca	tatttagatt	atattgtctg	acataagtg	ctattgttaa
13321	ggtttgacaa	ccattagcta	tgcagatatt	tgcattattg	aattaaattg	gatctgcata
13381	gtttgaatgg	aatgataccc	ttaagctttt	gccttcggca	tcaattaatg	tgtgctcagt
13441	ctctcttggg	gtctggcatc	acattatggt	ctctttatgt	gtaaatatat	ttaggagcaa
13501	atcaaatacc	actttcaatt	ttcaaataag	cttactcgg	ttataattaa	atttctannn
13561	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
13621	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
13681	tgactnncta	nagnatcccc	tgtcatttta	taggtgaatg	aagagagtac	ccttttaaaa
13741	aaattttgtg	tttcaaatag	aaaagttgag	caattattaa	taaggttgaa	cttttttttt
13801	cttttatgtc	ttggctaatt	ttgcccctt	ggagaatgct	ttttattttt	ccatagggac
13861	tcagtgaat	tgtctttttc	ttattgatta	taagaatgct	ttataaaaaca	aggctgttca
13921	cccagttgaa	tcacctttta	aatctctttg	caattaataa	atattacatt	gtataattaa
13981	atgtcagtca	tttccctttg	ggtttatttc	agcctttggt	atcatattaa	tagctctgct
14041	tttttacatt	tacatctgta	attagctctg	aatttagtgt	aaggtggaca	tagagactta
14101	gtcttttttt	ttttttgagt	aagtctcaat	ctgttgccca	ggctagagtg	cagtggcttg
14161	atctcagctc	actgcagcct	ccacctgctg	tgatcaagca	attctcctgc	ctcagcctta
14221	caagtaactg	ggattacagg	catgtgccag	cacatccagc	taatttttgt	attttttagta
14281	gaaatgggg	ttcacctgtg	tgccaggct	tgtcttgaa	tcctggcctc	aagtgatctg
14341	ccgcctnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
14401	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
14461	accccatttt	ggctctgagg	aaccccat	gatctaacca	actgtaaatt	cgtttgagtt
14521	aataccatcg	atcgtcattt	ttgaacggcg	tgaatcagtg	aacaagtacc	atttgccgtt
14581	cattttgaaa	acattcgcg	gctcgatttc	atcagttacc	gtgtttgaag	tgatcagcgg
14641	cttcattact	tttttcaatg	tgtaatcatt	atttaactct	atgataccga	gggcgcggtt
14701	cgctaactca	gcacgcggtt	ttttgagcgt	ctgctgaagc	ttctggcttt	ctttacggaa
14761	gaagttcgtg	ccgcgcgcgt	agtagcgttt	gttaataaaa	gattcttcgc	cttgggtatcc
14821	gtttttctgt	ccgtgtgttg	cttcgaatac	aaggtattta	tggcctttgt	cttcaacgta
14881	gtgaggtct	ctcagcgtat	ggttgctgcc	ggatgtataa	ttgccttcac	cgataaactg
14941	ctgaacgttc	tgatatgttt	ttccgtctcc	gtcaaaaatc	gttttgtgat	cttccactcc
15001	gttgattttg	agtgtgtcat	cagattttga	cacatttacc	tgcgctgttg	tcaggctttg

15061	tttgccgtaa	tgtttaccgg	aatagtcagt	gtagaataaaa	cggatttttc	cgtcagatgt
15121	aaagggtgca	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
15181	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnagctagac
15241	gttgccagact	cgcagtcgca	ctctanagaa	ccccctgggtg	ttccttttct	attgaagtaa
15301	ttcattaaaa	tttccattcc	atttcttatac	tttttcttct	tccttctgaa	actgtgtaga
15361	ttttcatgtg	aggtcaatta	atcatcttgc	ttttactcat	gtagtaagca	catttttctt
15421	tgcaataact	atcttctct	gtagaaggag	aatggtctaa	ttttaacaat	gggcaaatt
15481	cgtaactggg	gaactactcc	tacttattac	ttaattgggtg	attctctaac	tagagttaa
15541	aatatttgta	attctttcaa	agattattgc	ttaaagactg	aaaacactgt	acatgcctgg
15601	gtggggaata	ggaatatagt	gatattttaga	gaaatcttta	agaccagcat	aattcttta
15661	ataagctcga	gaccctggtc	cttggtcagt	tacaataaaa	gattaccatt	gggattttta
15721	atcaactaat	ttatgatgcc	tttagtttcc	tattatatta	aaatgttct	taatgtctta
15781	atttttagaa	gtattttata	gttaaataat	ttgagaagcc	tcaggagaat	aaaatcaa
15841	taatgatgaa	ttctcatggc	aaaggaaaaa	atgccctttt	tttcccactg	cagtttctt
15901	ttttaaaaaa	tannnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
15961	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnacactttt
16021	gctagaccg	aacacttctc	tanagnaacc	ccatacaaag	tggtgttttt	tactttttat
16081	tcaccaatag	caaacactaa	caggctgggtg	actcacactc	tcctgttaac	gccgctcca
16141	acaggcctct	ctctttttat	tcttactgct	atcattccag	ctaaacttct	gccaccttac
16201	acatgtatga	aactgacact	gtctcttctt	tctcaagtct	ctgcctcttc	caattcacct
16261	ttcattttgt	gccagagcaa	accacaaaaa	ggaaaatctg	attaacaat	gcttattaag
16321	caatatttga	tgccagggac	tattaaaaag	ctccaatatt	taacttgatg	ctactagcat
16381	tttagggaca	agataaaatta	acaaaataaa	tgttatgtta	aactgtgaga	actgctacag
16441	agaaaaaaat	ttttaaaagat	ggaggataga	aataggaaac	agggaggagg	gtataagaga
16501	attacaattt	taaacagaag	atctccctga	taaagaaaca	tgtgagcaga	gttctgagg
16561	aggtactgag	taagtcatat	agccacctgg	aagaacattc	atgcatcgta	ttatttcata
16621	actcaaacat	ttaaaacgaa	taccgaataa	aggagctctt	aaactattat	acaaacatct
16681	caaagcaaaa	gaaacttcca	tatcttcctc	attgcctgat	accaaccag	annnnnnnnn
16741	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
16801	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nacacttttg	cagaacccaa	cttctctang
16861	agaacccttt	tgtgttgctg	gacatagctc	ttgactggat	atttatactg	caaggaaact
16921	ctcccactct	gtggtttgca	ttttcatctc	ttaacgatat	cttttagggg	aaaatttctt
16981	gattttaaca	tagctcatag	gttttgatga	caggcctcat	aagtaagtta	cagggtata
17041	ctggcctcat	aagtgagttg	ggaagtattc	cctcatcttc	taggtttttt	tcttttctt
17101	ttcttttttt	tttttttttt	gagatggagt	cttattctat	tgctggggt	ggagtgcagt
17161	ggctcaatct	cggttcactg	caacctccac	ctcccgggtt	caagcgaccc	tctctccaca
17221	gccccctgag	tagctgagac	tacagggtgtg	caccaccatg	cttggcta	tttgtatttt
17281	ttagtagaga	cagggtttag	ccacgttggc	caggctggtc	tcgcactcct	gacatcaagt
17341	gatccacctg	cctcggcctc	tcaaagtact	gggattacag	gcgtgagcca	ctgcactctt
17401	ctctgttttt	taagagtttg	ttaagaattg	gtttgagttc	ttcaaatggt	tggtataatt
17461	tcaatggtga	ggccatctag	gcttgagctt	ttctttgtgt	atagnannnn	nnnnnnnnnn
17521	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
17581	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnggac	tacgtgncag	aaaaccacga	ctcactctan
17641	gaganaaaacc	ccctcctctt	tgctcaatat	gttttatattg	cgttggcaac	tgaattaaga
17701	aattgcattg	ccaataaacc	ggtccgtggg	gagaccacaa	cgccttctgt	ggacagtata
17761	tgcttggttac	ctattggcat	gctcgtagct	ctgtaaaaaa	aaggcaaagc	tcttaagatg
17821	tgacgttagg	ctgtgtatcc	ctttctcact	taaaagacac	aaagccagcc	ataagaagac
17881	agatcgatgc	aatgattcaa	tattgtctgc	acctaaggga	gtaagaaaaa	aacccttttg
17941	tttggttgcg	agtttcctgc	agggaagaat	cttgagtgat	gggggagtga	gaacacactg
18001	gaatacttat	ttgccggatg	acaccggctg	tctcaacgcc	catgcaatat	taactcttgg
18061	agtctccagg	agaaatgaca	aaaagtccgt	ttctcaccat	gacactcgga	gactcaccca
18121	gctttcttta	tgtgctgcat	cacagggtgta	agtaagactt	gtggaaagag	ggctgatgaa
18181	ggtgcgtgca	aggcatgtgg	ctgcaaaggc	caccttatct	acagcattca	gaccagcag
18241	aggctgacag	tggtgactc	actgtgtgtc	ctgcanggga	ggtgctgcag	atggaactca
18301	gcttctcagg	aatctgcagg	tcaccagca	tgcataaccn	nnnnnnnnnn	nnnnnnnnnn
18361	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
18421	nnnnnnnnnn	nnnnnnnnnt	gcttgttgct	gaaccgatct	gactctanag	naccctata
18481	ttatgtgtac	ctggctcatc	tagctgttat	gatataacta	aataagttag	tatacctaaa
18541	gcatttaaaa	cactacttag	cacataataa	ttagaagcat	tgatgataat	ggtggttatt
18601	atgtcagtgt	tagttaaaat	ttaacgtttg	tcctctgtgt	aaaccaagtg	ccatttggtg

18661	cccctccttg	gtaatacaaa	ataggagaat	agctctgtgc	tccttaatat	atctaaagac
18721	tttaataaga	aaacatacaa	agatgactgc	cagctgcctt	ctgcctttcc	atcttgtagt
18781	ttttgcttat	actgtattct	ccttgtagtg	tttgcttata	ctgcattctt	tgcatagaat
18841	gttaacagaa	gcacatcctt	ctcaaacaaa	acataataaa	aaaaacttgg	taagctattc
18901	attctgtatc	tggttacttt	cttttattct	aatcaatcta	ttctgttttc	atgtctgtgt
18961	tcccctagtt	tatatctgcc	aatgttttagg	actgacaaat	aataagtagt	ttttcattta
19021	aaattcctac	tgaactcgaa	atcattnagt	attgatcagc	tcaccctggc	caagaaattt
19081	tcctcattgt	tactccatgc	ggcccccccn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
19141	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
19201	nnnnnnnnna	ctacttttgc	tagaccctca	acttctctan	gagaaccctt	ggtattttaat
19261	atthgggtgt	ataatgaaga	tccccaaatt	caagtgacta	gatgtagtaa	aagcttattt
19321	cttgcttctg	aaaatctagt	gggtgtttga	atgacctctt	aatcagtggc	ttagggacac
19381	agactcattc	tgattagtga	cttctcttta	aatccactg	gaccttggc	atcaagccag
19441	ccagttagta	gagcctgtgg	agaagacata	ttgctcatac	ctacctcagt	gcataatgtg
19501	catacttcat	tttactccca	ctctcttgac	cagaactcat	tcataatatc	tcacttaact
19561	gcacaggaag	ctgcaagtag	agtcctcctt	atgcccagga	aattgaagtg	gatattgggtg
19621	agccctagca	gtctctaaca	cattcatgta	tatggtcata	aactatatgc	tatgaagtct
19681	gtgtcgaaat	gtttctttac	cctacctcta	aactgaattg	ggcaaaactg	aaaatgtttg
19741	aacctgatgc	cttctggaca	tagataacct	atagaagtct	aaatgtaatc	ttgagttact
19801	acttatacta	tgtatatatta	ttgaaatttt	tgtagttttt	ataatttgaa	ttacaaaata
19861	aattacaaaa	taaattttgt	aatttgaaaa	tgaatatatgc	nnnnnnnnnn	nnnnnnnnnn
19921	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
19981	nnnnnnnnnn	nnnnnnnnnn	acacttcttg	ctgaacctat	tttactctan	gagacccctt
20041	ctttaaatgt	tttttgacta	aatgaagcct	catattgccc	taatttggtg	ggaaagacag
20101	ttttgtattg	gcctttcaag	cctacaattc	ttggtgttca	gtgtgggaga	atagagcacc
20161	agtttttcat	taaaatgtgg	cattaaaatg	tctaatagaa	aataaattca	cttgtttcaa
20221	catcagaagg	atgcacaaaa	agaatatacg	gtgaaaagtt	ccccaccca	tgtatcccca
20281	gccacttagt	ttcctttccc	agaggcagag	ccattgttag	tagtttctaa	acgtactttt
20341	ccagagatat	tttattcata	tacaagtaag	ggcatgtaaa	tatgcacgtg	cacacatgca
20401	tgcacataca	cacgcatatt	ttgtttccct	ctctctcctt	atgccaatgg	catattttag
20461	tcacccttct	gatttcaaaa	attaaaacac	acacacacac	acacacacac	acacacacac
20521	acacacaaac	atataattttg	ttttcttttt	tttttcttga	gatagagtct	tgctctgtca
20581	tccaggctgg	agtgtagtgg	cgtgatctct	gctcactgca	acctccacct	ttcttgttca
20641	agtgattttc	ctggcttagc	ctcctgagta	gctgggacta	caggcgtatg	ccnnnnnnnn
20701	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
20761	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnngtgcact	tcttgctaga	tcgcagcttg
20821	actctanaga	accccaattt	ttgtagtttt	tataatttga	attacaaaat	aaattacaaa
20881	ataaattttg	taatttgaaa	atgaaatatg	taaacagttt	atttgtagat	gtgtagagtt
20941	gaaatatatt	ttcctccctt	cattttcagc	ctctgcataa	gagcctagat	ccaagcaacc
21001	tggaacatct	cataacacca	ttggttacta	ttggtcataa	tgctctcctt	gcacctgatc
21061	aatttgctgc	tcctttgaaa	tcctttggtg	ctactttcat	tgtgaaagat	cttctcatga
21121	atgatcggtg	aatttatatt	tttttagattc	atgttctttg	taatgttaac	ttttgcaatt
21181	ttatccaaca	aggaaagaag	tgaatgata	tcaataatac	ttccatttga	ctgccttgat
21241	tttaggtttt	cctgtttaaa	aaaaataagc	cttgtctcca	ttataagcaa	aacctcttg
21301	aaacatcttt	ttagttgaga	ataggcatat	gacctcaaat	acattccttg	ataaatataa
21361	aatataacct	gatgctatct	atttcagatt	tacaggggtt	taatttttga	caggacatct
21421	attttttaaaa	atatgcagat	tggtaaacca	gctcatagta	tgggnnnnnn	nnnnnnnnnn
21481	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
21541	nnnnnnnnnn	nnnnnnnnnn	nnnaatctt	cttgctaaac	gcattttgct	ctanagaacc
21601	ccatgttttc	accattttga	ggtggagcat	catacaactt	gtcacctatg	cttctataaa
21661	acagagactt	ttaaaccattt	aaagaatact	atataaaact	gccacagtgc	tttgtagata
21721	gtaaacatgt	aacaaatatg	tactgaattc	gtttgattct	cacaattaat	tgtgtcataa
21781	tcacccctga	ttcaatctct	aataacatat	ccaaaactga	attcctgggc	tttccaccgg
21841	tgcccacctc	caaaccgggt	ccaacctcag	ctttcccat	ctccatctgt	agcaatttca
21901	ttatttttagt	tattcatgcy	caagccatga	agtcgattcc	atttccctca	cacatccttt
21961	ctaactctgtc	aggaaatcct	gttgactctt	cttgcaaaat	atatccagga	tctaaccacc
22021	tccactgcta	cttccatcca	agccactatt	atatcttacc	tgggtttatca	gtttgctagg
22081	gccaccataa	aaactactac	aaactggata	tcttaacaga	aatttattag	ctcacaaattc
22141	tggaggccgg	aaattctgga	gaccaaaatt	aaggtgctgg	tagcgttggt	tccttctgag
22201	ggcaggattt	attccaggtt	tctctctggc	ttgtagatag	ctatannnnn	nnnnnnnnnn

22261	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
22321	nnnnnnnnnn	nnnnnnnnnn	nnnnnggcta	cttttgctga	accccagact	tgctgctgan
22381	agnaaccccc	ttggcagnat	gtttatttcc	atttagaaaag	attattcttt	tttgttttgt
22441	tttgttttgt	tttttgagac	acagtctcac	tttgtcgccc	aggctggagt	gcagtggcac
22501	gatctgggct	cactgcaacc	tccgcttccc	gggttcaggt	gattctcttg	catcacccct
22561	ctgggtaact	gggaccacag	gcattgtgcca	gcattgcgtag	ccaatttttg	tattttttaa
22621	aggcggggtt	tcaccatgtt	ggccaggctg	gtctcaaacg	cctgatgtca	ggtgatctgc
22681	ctgccttgge	ctcccaaagt	gctgggatta	caggcttgag	ccatactgcc	cgctgtgaaa
22741	gaattattct	aaaattttat	tgaattctaa	ctatgtatac	tgtaagttac	gatgtttatt
22801	ggctgtttaa	aattttcaca	tgagcaaaat	taagaattga	tcacttagct	tggaacaat
22861	cttgttatct	ggatttacag	aagaaaatcc	ttgaattgga	aaaggtatta	tgtcttctta
22921	gttctcactt	actgcagaaa	ttcctttatt	attaatgaga	gctggtcacc	acgttgaaaa
22981	tgctttaagc	agagagctca	tgatgtattt	catttgttgg	aaaagtcaac	taggtcactc
23041	ataccacann	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
23101	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
23161	gactaggtct	gaggggtgat	ccgttgngag	ttacgcttcc	ctctgggtga	agagaactac
23221	acttctcaaa	agagacattt	ctggccaaca	aacatatgaa	aaaaagccca	tcactactgg
23281	tcattagaga	aagcaaatca	aaaccacaat	gagataccat	ctcacatcag	ttagaattgc
23341	gatcattaaa	aagtcaggaa	acacttttgc	tggagaggat	gtggacaaat	aggaacgggn
23401	gtacactggt	ggggggggcg	tggatcccg	ttgcccggga	aatgaggggt	gcccggggag
23461	tagattattt	ggaccgggac	ggactgcagg	cagactgccc	gcccggggcg	ttgtggctcg
23521	ggggggggcc	cccctgttcc	cggtgcgcgt	ccgttccctg	cgcttatgct	acggcggtgc
23581	atttctaccg	tgtgcggccg	gtcccataatg	tgacccttcc	cgctcgcccg	gggttccgta
23641	cccagccccc	cgcgcgcggg	tcttcccccg	tccccccga	ggtgcgaggg	caggtccctt
23701	ttgcgcggcc	ccctcggttt	agcccttagt	gggtgtccgg	tgccatgccc	cccccgcccc
23761	ccccgcggcg	tgcttttcgc	cccgtgtctc	acatagcatg	ggggccggcc	agtcctctgg
23821	cttcgcggcc	cgcggggccc	ggcgcggggg	ttcccccnnn	nnnnnnnnnn	nnnnnnnnnn
23881	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
23941	nnnnnnnnnn	nnnnnnngac	acatttgcta	gcacactagt	acttctctna	nagnatcccc
24001	ctcctgggtc	agctacgctg	agttctaaac	tcacatatcc	aatnntgtct	actcaacnag
24061	cttcacttga	aattctaaata	ggcacctgat	ctttaatgta	gccaaaacag	ttatttgaac
24121	actcttctta	tttccacttc	cccaaaacac	atacatataa	acagagcctg	ttcttccotca
24181	aggcttctgt	atctcagtaa	acgggtaccac	catctctcca	gttattcagg	taaaaagccc
24241	aggatacacc	cttaaategt	ccctttccct	aactgttcac	atctaataca	ttagaagat
24301	catctgatct	atctaacatc	aaaataaate	tgagggtggt	cgaaaggcag	tgagttatct
24361	tcaactgatt	gctcatggtt	agtttagacat	caaactcctt	gttttaactc	actcctccct
24421	tctcactact	gcatttcaact	actcttaaaa	aaaaaatctg	aaatttgact	atttaccatc
24481	ttcttccctc	ttcatccctg	ccatagatta	tgattcgctg	actcctaatt	gctttttcta
24541	cctcaattct	agctcctcta	tagtacatct	tgacaaaatc	atcagggcaa	ttttttttaa
24601	aaaccacaaa	tgagattgct	agactccctg	gcaaactctt	cagttgcact	taaaatcaaa
24661	tccaaaactcc	ttaccatggt	ttacanannnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
24721	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
24781	nnnnnaggcc	tagcgttnca	tgagtgtac	gatgcacggg	nagnagaaaa	cctctgtctc
24841	cctgancttt	gttgaaattg	tttctcagct	taaggagtct	tttggctgag	acgacagggt
24901	tttctgaata	tacaatcgtg	ttatctgcaa	acagtgacta	tttgacttcc	tctcttctta
24961	cttgaatact	ctttgtttct	ttctcttgcc	tgattgccc	ggccagaact	tccaatacta
25021	tggtgaatag	gagtgggtgag	agagggcacc	cttgtcttgt	gcgggttttc	aaaggggaatg
25081	cttcagctt	ttgccattc	agtataatat	taaagaatgt	tttacctttt	tctgtcttgc
25141	ctgtttttct	gtgtttttgt	tgggtctctc	attctccatt	tttaggcctt	tacatgttag
25201	gaatatattt	cttttaatga	tacttcacct	ttggtatctt	ttgtgagact	ctactcttag
25261	tgtgataagc	actgggttgg	tataatgcac	tgaggaaaat	gtgaaggccc	ctacgttccc
25321	ttaggcgccc	tgttaaaagt	agttgacctt	gttccgcgaa	gtgattgttt	gccaatttta
25381	agtagttgtg	ggtttccata	acctcctccc	cccattacct	gtctccctct	gatccgttat
25441	aattatggta	aaaatatata	tatgtogaac	cccagttttc	cctcgcccaa	atgcaccccn
25501	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
25561	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
25621	gtttctactc	tanagnatcc	ccccaaactcc	tattagcgac	agtagtagaa	gatgggtatcg
25681	gggtctttgt	tgtgatcttc	tcttttatta	aacaagtttg	cgtccaaata	gctttttttt
25741	tttttttttt	tttttttttg	agataaagtc	cgatgggtga	cttgtcagtt	gggagctaaa
25801	gggttagagga	caggtccctt	tcaacatttg	gacaggccac	tgactaaaac	ctgatggtag

25861	ctgtcttttt	ctttgagttg	ctgcaaactt	gtgggtttggt	aacaaggaga	tttgtgggaa
25921	ctaattgtct	attttaaagt	tcaggtaaca	aatataagaa	tatccaggct	gtaactggag
25981	ttcttacatt	tgcttagcaa	cagaactgca	attcagacta	atattaaata	gagatttttc
26041	ctattggaat	actgcttttc	ctagtataac	ttttttccct	tttttagaga	attatttata
26101	aaggcaataa	acaggcacag	attgccttgc	tggtaataac	aattttatta	aaatacttga
26161	aaggatagtt	cctataaaca	tggtataata	caagtcttta	gagatggtca	cttttttatt
26221	ttattttaagc	tatagccata	acagggtttat	aatgggcttt	aaaatagctt	gacatttcat
26281	ctttaatgag	taattttccct	ttnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
26341	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
26401	nnggcaacta	tctgtgctga	acctcgcagc	cttactctga	naggaacccc	ctntttgtac
26461	attttacttt	ttcattaatt	agtcttttat	agtattttaat	agctccatag	cataccactg
26521	aatacatgta	ctataatttg	ttgagtagtc	cactaagtta	cacattttaga	ttttttccta
26581	ctttttgatt	actattaaga	agaatataat	taatattttt	acacagtctt	gattctttct
26641	ttaggataaaa	ttctaagaag	tacaattgca	tgggctaaca	gctcattttt	aaagccttta
26701	atatgctttg	gcaaattggt	ctccagaact	gaattaattt	tatgctaata	accagtaatg
26761	agagtgacca	ttttctcaca	cacaacaata	ttggtcagta	tcattgctgg	aaacaaaacc
26821	caagtaaaaa	caaagaaaaa	cacaattatt	ttgatagtat	ttatatttgt	tagtcaactg
26881	tcatctgacc	atttttttct	ccttgggatt	attggccatt	catatttctt	ttgtaagttg
26941	tgttttataat	atttgcctag	ttttatagaa	tggtatttat	cttactgatt	tgtgagtttg
27001	tcatatggca	tttgctaagt	tgagataaat	tttctgtcag	ttttagttaa	ccctttcttt
27061	ttgtttatta	tgttctttac	ataccatagt	ttctatcttt	aggtaagcnn	nnnnnnnnnn
27121	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
27181	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnna	ctatttgctg	aaatcgagtc	ctgctgctna
27241	nagnataccc	ctcaaaaaana	aaatgccgat	acttcattgg	catttttctt	tattttctcaa
27301	caagatgggtg	aattgactag	tgggtagatc	cacaggacgg	gtgtgggtcg	catgatcgcg
27361	tagtcgatag	tggtccaag	tagcgaagcg	agcaggactg	ggcggcgccg	aaagcggctg
27421	gacagtgtct	cgagaacggg	tgcgcataga	aattgcatca	acgcatatag	cgctagcagc
27481	acgccatagt	gactggcgat	gctgtcggaa	tggacgatat	cccgcgaagag	gcccggcagt
27541	accggcataa	ccaagcctat	gcctacagca	tccagggtga	cggtgccgag	gatgacgatg
27601	agcgcattgt	tagatttcat	acacgggtgc	tgactgcgtt	agcaatttaa	ctgtgataaa
27661	ctaccgcatt	aaagcttatc	gatgataagc	tgtcaaacat	gagaattgat	ccggaaccct
27721	taatataact	tcgtataatg	tatgctatac	gaagtattta	ggtccctcga	ctatagggtc
27781	accgtcgaca	gcgacacact	tgcatcggat	gcagcccggg	taacgtgccg	gcacggcctg
27841	ggtaaccagg	tattttgtcc	acataaccgt	gcgcaaaatg	ttgtggataa	gcaggacaca
27901	gcagcaatcc	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
27961	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	aaacctacgg
28021	tnctgaatgc	cagatcncgc	tanagataaa	ctttaacctt	acccccaccc	cgtgctctct
28081	gaaacatgtg	ctgtgtcaac	tcagggttaa	atggattaag	ggcgggtgaa	gatgtgcttt
28141	gttaaacaga	tgcttgaagg	cagcatgtct	attaagagtc	atcaccactc	cctaattctca
28201	agtaccagg	gacacaaaca	ctgttttagg	ccgcagggtc	ctctgcctag	gaaaaccaga
28261	gacctttggt	cacatgagta	cctgggtgcc	ttatgtcacc	tgttgacctg	tagccactca
28321	cagatacccc	acgtgaagcc	gccacacccc	gcccagagccg	tgaccccgtg	ccggtgctgg
28381	tgtcccgtctg	tccccccacc	gtccccccgt	cccgtcccgt	ccccagcccc	tttgccgtcg
28441	ccgtagccgt	tttagcacat	cccagcacc	cccggtagtc	cccctgttcc	gctgccttg
28501	gcccccttac	tctccccgcg	cccccccgct	gtgccccgtt	cccccgctat	atcgaccccg
28561	gccccctgag	ccgccccccc	tatcggtttc	cccccttggg	acatccgggt	ccctctcccc
28621	ccgcccccg	cccccttcc	tttcgcccc	tcttccattt	cgccataaccg	ccccctgccc
28681	ccctgatccg	cccccttctt	tcccccccc	ntttttcccg	nnnnnnnnnn	nnnnnnnnnn
28741	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
28801	nnnnnnnnnn	nnnnnnnnnn	gcacttcttg	ctgacctagt	tttcttagag	acccaataa
28861	ttgaattgggt	ttgaaatgta	atattttaaa	cttacacatt	tttttttggg	tgtgatgggtg
28921	atggcggtttt	tgatgggtctt	cctcgctcct	tctgtgggtg	ggactgtgtg	gattcaattg
28981	cactagattc	aggagattct	gctctgcttt	ggggaagacc	caaatgtgga	tagcatcaaa
29041	ttatgattta	taatatagta	acgaaacaga	tgcagtttag	agttacacat	gcttacctct
29101	gctgtgctct	ccttgacact	cggtgctgtt	tgcttttggg	cttctgttcc	gtattttcac
29161	tttgtctttc	ttcttcttcc	tctctctctg	gtgctggagg	tccagatttt	tttttgcttc
29221	ctttttttaga	agttttcatt	gttggtctct	cttttgggtg	acctccacca	agaggttttg
29281	gtggctggcc	tctttttacc	ttttttggcg	gactattctg	ttcatcttca	tttttctaata
29341	tatcttcttt	gagcctcttt	tcttcaggcc	actgctgttc	atcagattct	gaagccgtat
29401	ggcctctntt	ccgacttcgc	tgactgcctt	taggtttctg	ttcctgtacc	aacttagttc

29461	agtcacatccat	acctaatttc	tcctnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
29521	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
29581	nnnngcacat	acgtgctgaa	ctgcacgcat	gagtgtngng	gaaaacactc	taaaagcatt
29641	tatacacttg	caagtataaa	agtatcagtg	ggaacttgaa	nacacagcaa	tagaaactat
29701	caaaaataaaa	tgagagagaa	aaaagggcaa	ngaaaaaaat	agtaagacag	cagctgtgat
29761	ctataagaca	gtatcaagta	gtcaaaaata	ctntgtaatn	ntggagcccc	agtaaaacaa
29821	caaggcagat	ggagagaata	gataaaaaaa	aaaagtattc	aaaataataa	tggccaattt
29881	ttttccaaat	ctgattaaaa	taaccttaga	gatccaagaa	tctcaataaa	cctcaagcag
29941	aaaaaagata	aaccatttca	aggcatatta	ttattcattt	ggataaaaaa	tagtgaaaaa
30001	cagaaaaatg	ttaaaagcag	tgcaaaaaaca	gtggaacaag	gattgataga	acacttttga
30061	tcagaatctg	caatgcggaa	catgtgcccg	tgcccccttt	tgatcgtttg	ctgtcccccg
30121	ccccccgggt	cctccgctgg	tgcccccggt	agtcatgccg	gggctgagtg	tcgcccgcgc
30181	tttattttgt	tcccgggtgt	gtcccgtgtg	cctgcccccg	gccgcctgtc	gccccgagac
30241	cggtcgccgc	agggaccgtt	ggcaggttgg	cgcgcgcgtg	gcggcaggag	tgtcggcccc
30301	gttgtgcgca	cgaccgagtg	ttgcgggttn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
30361	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
30421	nnnnnnnnnga	cctagtgcag	aatctcattt	ggctctnana	gaaaacactg	taaattgggt
30481	caaaacaacc	ttattcacat	acttgtcagg	aaacataact	gctcaccatt	tttaatccat
30541	taattttgtct	tgttttatta	tatatttgat	atgtacaaaag	gaatatatgt	aacatttgta
30601	taacatgttt	cataatacaa	tgaacaccta	tgaactctta	gttccttaga	gttcttagag
30661	ttttattact	cttcaatcct	ttcattctcc	tccttgctcc	taatatctac	tctcttgaat
30721	cttggggttaa	ctgtaacag	caacaaaatt	taccacatac	ttgtttagtt	ttgtctgttc
30781	ttgaggtttc	taagcattat	atcataccct	gtatagtctt	ggaggatttg	cttttatcac
30841	tccacatttg	tttctaagaa	taatcttcgt	tactaaatct	aactgtgggt	cattcatttg
30901	ttactctata	atttctactgt	ggaaaatatc	tcgtaattta	tttattcttc	tagtagattg
30961	tggaatatatt	ccatttcttt	actctttcgg	acagtgtctgc	tatgaatgtt	tctattgaag
31021	ccccctgcta	catatgagca	agaatttctc	taggttttat	tcctacgagt	gaagttgatg
31081	aatcacggta	tatagacata	tttaactttt	cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
31141	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
31201	nnnnnnnnnn	nacacttttg	ctgaactcca	gtactgtact	ctanagnatc	cccctctact
31261	tcaagctctg	tcactaactt	attgtgtaac	ctggaaaaat	cctaactgct	gtgtctcca
31321	gtttcaacat	ttctaaaata	ggattctttt	atttgtattg	aatatgtatt	gactgtcttc
31381	tagtagaccg	tgtgctggac	cctgggattg	caacagagag	caaacccac	taagccatat
31441	tctaattggag	gggcagacaa	taagcaaaat	gtgaacatct	attatagtgt	gtaccagaag
31501	cgggtgagtg	ccatgcagaa	aagtaaagca	taatacagta	gtaaggagtg	ctggggtagg
31561	ggatgacatt	ttaaatcaga	aagtcaggaa	ggcctgtgtg	agtcagaaac	atctgagtag
31621	acttgatgga	gatgagaaca	agactcacct	acacctgggt	gaataacatt	ccaggaagag
31681	gaaatactac	atacaagcac	tatgacaagg	aaggaaactt	gccttgtagc	ttagaggggc
31741	agcagaagaag	cctggggggc	gaagcacagc	ctgtgagggg	agagtaggag	gcgataaagt
31801	cagagatgtg	acggcgacca	ggtcatacgg	cctgtgagac	cactgtaagg	accttggtt
31861	tcattccaat	gttgatgcat	tgnggaattt	taagcaaaaag	agacatagga	cctgaattcc
31921	atttttaaaag	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
31981	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nagaactaga
32041	cgttgctgac	ntccccgtgc	ggctnctnan	agaatcccct	gcacatatac	ccattttaat
32101	taagcatctg	gtgactttcc	ttttaaggta	cggcggcgaa	gtgctaaaag	ggaacggcga
32161	tgaacaaatg	taattaataa	ctttctctgt	gaaagctttg	gaaaaatctt	tttttttttt
32221	tttttttttt	gggcaagctt	gaggctgaat	aaagcctttg	atgcacaaaa	tgggactgct
32281	gaagagtggg	cagttggacc	ttacttttgt	gaccccatat	atttgtgggt	acatgcttta
32341	gccatacaca	tggttaacatt	gactatggag	tcttgtgaaa	gtgtaatgtg	cgatggctat
32401	gtagacataa	agaagaaact	tggaaatatc	ttttttcttt	tttttaatgg	ttctgatttc
32461	tgaagtgtct	gtatagcttt	tatctgcggc	tttaaaactga	cagtacccca	ctgtttattg
32521	gatctattga	tttgaaaaga	atttggttagg	atagatctta	agcagtaatc	tgtcactgtt
32581	tgtattttgtc	ttgtgtggaa	ttttgctgtg	aaaaaaaatt	tgttgtcaac	aatcgggtgcc
32641	ccctgtgttg	atggcacctt	ttttctagag	aaggggtttc	acccccccat	tcgctcccc
32701	gtgcttcccc	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
32761	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	acactttttg
32821	ctagccccc	gtcttnactc	tanagnatcc	cccnacagag	ttaagactaa	tgctggcact
32881	ataatccagt	aaacaaaata	catccttatt	ctctcaaaca	gaaattaacc	tgggttgtgg
32941	atatagatga	ggtcagaaga	ttgactgtca	ccaagtacct	caaggcattt	ccaagggctc
33001	tactctttat	tattattttt	ttctccagct	ttattgaggt	atacttgaca	aataaaaagt

33061	gtatacat	gaagtgtaca	atgtgatatt	ttgatataca	tatatattgt	gaaatgggtg
33121	ccacaatcaa	gctaacatat	ccacacctca	cctagttact	cgtgtgtgtg	tgtgggtgaca
33181	acacttgatc	tctgctctct	tagcaaattt	cagtatatat	aataacagta	ttattaacca
33241	tagtcaccat	gctgtttaga	aggtatctat	cccacataac	tgaaactttg	taccctttga
33301	ccagtgtctt	cctactcctc	tctcatccat	cccctggcaa	ccatcattct	actctctgct
33361	tctctcagtt	tgacctccta	gattccacat	ataagcgaga	tcatacagta	ttcttctttc
33421	tgggtctggc	ttgtttcatt	taacgtaatg	tcctccagct	tcattccatgt	tgttgcaaat
33481	ggcaggattt	ccttctcttt	tgnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
33541	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
33601	nnagactagc	gtgcaagagt	ctnagctcgg	ctctanagaa	tacccgtaaa	ccagagacct
33661	ttgttcactt	gtttatctgc	tgaccttccc	tccactattg	tcctatgacc	ctgccaaatc
33721	cccctctgcg	agaaacaccc	agaatgatc	aataaaaaaa	aaaaaaaata	aataaatcaa
33781	taaattaaaa	aaaaaaaaaa	agcaccaatg	taagttgtgt	cttgaggagg	atatttcaaa
33841	tctataagaa	accaggaatt	ctaattcttt	taattgttca	gagtatagaa	aagaaagcac
33901	tcattttctt	tacagccagg	ataactaaca	tgaatactaa	aacctagtaa	aagatatcct
33961	ccccacccac	cccccaaaa	aaaaacgcta	agacaaatct	tactaatatg	tgatttttgt
34021	ataagcaaga	tagtagcaaa	aagagaattt	aaataataaa	gcgagaatat	attatggcca
34081	aacatggtta	attttaggag	tatacagaga	gcgttggtac	aatattagta	agcctactat
34141	ttatgtttac	aattttattat	ataagtaggg	aaaaggggat	acaatagatc	atttttgata
34201	gatgctgaaa	gggcttctga	attttttttt	ttttgagacg	gagtctcgct	ctgtcgccca
34261	ggctggaggg	cagtggcatg	atctcggctc	actgcaagct	annnnnnnnn	nnnnnnnnnn
34321	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
34381	nnnnnnnnnn	nnnnnnnnnn	ngaaactaacg	ttgctgaaat	cccagctcna	ctctnanaga
34441	atccccaatg	gcaacaggta	tatggggttt	aaaagaaata	attgaagagt	gattgaaact
34501	actaaaaaaa	aaaatcaaag	cctcactttt	tttttttgag	actgagtctt	gctctatcgc
34561	ccaggctggg	gtgcagtggc	gcgatctctc	ggcttactgc	aacctctgcc	tcccagggtc
34621	aagcaattct	cctgtctcag	cctcccgggt	agctgcgatt	acagggtgtgc	accaccacac
34681	ccggctaatt	tttgtatttt	ttagtagagg	tggggtttca	ccatgttggc	tgggctgggtc
34741	tcaaactcct	gacctcaagc	aatctgcccg	cctcggcctc	ccaaagtgca	gggattacag
34801	gcatgagcca	ccgtgcccag	cctcaagcct	cactctttca	aacaaaagca	aggattgctt
34861	tattactctt	atacttttat	cactttatct	agcagtgaaa	acctctgaaa	cctgccatca
34921	tatttctagct	aggggagaaa	gaaagaatct	tctatagaat	gagataccgc	aagccctcat
34981	ctgtaacaac	agtgagtaga	agttaaaaag	caagaaatga	ctgactctaa	ggctactggg
35041	aaacaggaga	tgggaagcca	atatgagcct	aagaaggaaa	cgtatacggt	ggcaggatat
35101	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
35161	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	gacatacgty	ctgacgtcga
35221	cctcggctct	anagataccc	ctgtctccca	aanataactg	ctcaaaataa	cacaaaaggc
35281	gcaaatacaa	agctattcac	tgaaagactg	taatactgaa	aacaacccaa	tgtccataaa
35341	taatgaactg	gttaaataaa	ttatgatata	gccacacaat	aaatggggat	ctctctatat
35401	ttatatgaga	taattgccag	tgcatgccat	tagagagaaa	aaaggtgaag	aaaagtgtat
35461	atagttttatc	ttattcattt	tatcaaattg	gcatacatat	gcagacagg	atgtgcacgg
35521	aagggggctn	tnnttttnann	tnngccttaa	nttgtgagt	gtgctcagtt	tggagcacgc
35581	ctgcccgcct	ccattccctc	ttcctgcgct	tttctgccc	cgtgtccttc	cccacgtcca
35641	gtttcgccgt	tccgcccgcct	tccgtattgc	tccctttgcc	gcacatcctt	gtcttccctac
35701	tcctgctcgc	cccccccctt	cctaccttgt	ccccccctgt	tgtcctccgg	catgtctatg
35761	gccggccacc	attgttgttc	cccagggtgcg	gcgccctttc	ccactttccc	taacaccgtc
35821	cccccttgtc	cttcgcccc	gtcttcccgt	cccagtgtgc	cctccccttc	ctcttcgacc
35881	gacccccctt	ttcccaccac	ccatttttat	ccccnnnnnn	nnnnnnnnnn	nnnnnnnnnn
35941	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
36001	nnnnnnnnnn	nnnnngacat	ggtgcagaat	ctgcatctgc	tcntagagna	acccatcgca
36061	attcgagtga	accccaatta	tgaacatagt	aatgtagtc	ttcaagttca	catataattc
36121	tgaaatccca	ttttatgtct	tggtgagata	tttaaataac	ttaatatagt	taagaatcta
36181	aaaatatatt	aatcctccta	aatacttctg	tatgaccttt	tcctgtttct	gtaatatgaaa
36241	aaaatctgca	aggaaagaga	accacattat	agtacaacag	ataaacatga	tcacttatag
36301	aagataaatg	gaatgttttt	ataagttttac	ataataagca	cctgagggga	agtggaataa
36361	ctgctagtgc	agcacttcgc	aatctgtgtt	ccaaagacca	ctagtagttc	agtggaagtt
36421	aacagtgttc	aggaaaaggt	attttttcaa	ataattttgg	aaagtgttat	tttacagccc
36481	cttcataatg	gatgggcac	ataataaagg	ctacaataag	gcctattatt	tcagaaaaaa
36541	aaagcaaaaag	ggggtaggga	aaagaaggga	taaagaaaac	tctttcaatt	tagcattcta
36601	taaacatata	tgacaatgaa	accctttttt	ctcacaaat	tattaacaac	tgcaggatac

36661	cagcattcct	tgaaaacaga	atctgggaaa	tattgaactc	actggaagan	nnnnnnnnnn
36721	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
36781	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnng	aactagtgc	gactccagat	ggctctanga
36841	gaacactatg	atgtgtttta	atgtactttt	gtaatactga	aataatctgt	attacagcgt
36901	gaaataacta	agaagttggg	caaccccaaa	cagcctacaa	atcctttcct	ggaaatgatc
36961	aagttttctt	tggagaggat	agcacctgtg	cacatagata	ccgaatctat	cagggtattg
37021	tatataaaat	actaagttct	cattgctgtt	cattaatgtg	tacattcagg	aatcagtc
37081	tggtgagtg	ctgctatatc	tagctacaca	tctgtgtgta	ttgtgtactc	ctctgtatat
37141	actcctctgc	atatacctgt	atgtgtatat	gcacacacac	acacacagac	acacacatat
37201	atataatatac	actcatctgt	acctgtatat	ttgtgtttat	tctgtttatt	tatcttaatt
37261	ttaaggattg	gttcaaattg	tgtgagaaat	tgccaatcta	agcctaattt	tatctgttca
37321	cggatatacca	gacactcttt	gttttcgcct	acttgtatag	ttaaggtaat	ataaataaac
37381	tctattttata	ccctcgggaa	actaggcgaa	tctcccgccc	ccgcacagcg	gaccaccagg
37441	ccatgtcctc	cggaaactgt	tcaggagacc	ccaaatcagc	ttaattcaac	cccctgtttc
37501	cgaacccaag	ttatcccgnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
37561	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
37621	actatgtgct	gcaacctgat	gatgtangga	acaaccagac	acaaacaatc	tgctgttctc
37681	acttttaaag	ggaggaacat	caaacacaca	tgggtataca	cnatggaaac	aacagacaat
37741	gcggactact	gggggagctg	gagaaaggag	gatgtgggct	gtaaaacgac	ctattgggta
37801	ctatactcac	tatgtggatg	caatataccc	atgtaacaaa	cctgcacatg	taccctctgt
37861	atctaaaaata	aaagttgaat	ttaaaaaaaa	caaaaaagca	aaaacacatc	tttaaaaaaa
37921	cgaagtttat	atatgtacgt	ggtactatac	aaataaattc	aaatgatacc	aaaaagtata
37981	aaataacctc	cctttctata	ctaggttaac	ataaatactg	ttaagatttt	gcataattctt
38041	caaaaaaact	ttctatgcat	ataaaaacaa	atgtgaacat	taatctgttt	tcaaaactta
38101	aagattccat	aatgctgctt	tatatacttt	gcctttttca	ctcaataatg	tattgtggaa
38161	actcttccat	actaactcag	acctcatcat	gtttaatgac	tatgtattcc	attgtatcat
38221	acaccataat	acaacagttt	tacaattgct	acacactgag	gttggaact	agtttatgtg
38281	tactaacaca	gattaccccc	aacaagctaa	agaatgaaag	actaagacag	nnnnnnnnnn
38341	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
38401	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
38461	nggaataccc	caccgctcan	aaatctgatg	gcaccctaca	ggaacatgac	ggatctcgcg
38521	agatccatgt	tgctaaatat	gctgaaatat	tcggattgac	ctctgcggaa	gccagtaagg
38581	atatacggca	ggcattgaag	agtttcgcgg	ggaaggaagt	ggttttttat	cgccctgaag
38641	aggatgccgg	cgatgaaaaa	ggctatgaat	cttttccttg	gtttatcaaa	cgtgcgcaca
38701	gtccatccag	agggctttac	agtgtacata	tcaaccata	tctcattccc	ttctttatcg
38761	ggttacagaa	ccggtttacg	cagtttcggc	ttagtgaac	aaaagaaatc	accaatccgt
38821	atgccatgcg	tttatacgaa	tcctgtgtgc	agtatcgtaa	gccggatggc	tcaggcatcg
38881	tctctctgaa	aatcgactgg	atcatagagc	gttaccagct	gcctcaaagt	taccagcgta
38941	tgctgtactt	ccgcgcgcgc	ttcctgcagg	tctgtgttaa	tgagatcaac	agcagaactc
39001	caatgcgcct	ctcatacatt	gagaaaaaga	aaggccgcca	gacgactcat	atcgtatttt
39061	ccttcgcgca	tatcacttcc	atgacgacag	gatagtctga	gggttatctn	nnnnnnnnnn
39121	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
39181	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnna	actttttgct	gaatctnatc	ttactctang
39241	agaaccccaa	gtcaggagat	cgagaccatc	ctggctaaca	cgngtgaaac	ccagtctcta
39301	ttaaaaatac	aaaaaattag	ccaggcgtgg	tggcaggcac	ctgtagtccc	agctacatgg
39361	gaggctgagg	caggaggatg	gcctgaaccc	gggaggtgga	gcttgacagt	agctgagatg
39421	gcaccactgc	attccagcct	gggcgacaga	gcgagactct	tgtctccaaa	aaaaaaaaaa
39481	aaaaaaaaag	atttaaaggc	cctgattttt	tattctttgt	gttgaaatat	ttgcaatggg
39541	atgaccactt	tctttacatt	ttaatgaaaa	ctgtcagtga	ttactgtaaa	tttaatat
39601	tggtctcatt	tttctgattt	gttaaaagag	gatataagtt	gtgataagtg	tttgtaagc
39661	ttcagatctc	cagtctgatt	tcttgattct	tatttagctc	aataattacg	ttcagtcagt
39721	caagttaagc	caactttttt	cttttaaaata	caatggaaaag	tttttagaca	gtaggaaaaga
39781	atatcagttt	aaattatgca	acttaattta	aagtttaata	ttttaaaaca	aaaatacttt
39841	tactggatat	atattttatta	tatcttttaa	ttaagacttg	gatttagctgg	ggatgtattt
39901	tattgggttt	tannnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
39961	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
40021	ttgctgacac	cgnatctttg	actctanaga	atccctctc	ctattcaact	attcttcatt
40081	tacaacaata	gtgacaccag	taattatgag	aggctgtcat	atatcaaagc	agctcaggtg
40141	ggatcaacat	acattgggtt	tgaaatgttt	ccttagctcg	gttctcccaa	acacataatt
40201	tctgttttcag	cagtacaaag	gcatgtttta	aaatctataa	cataaagaat	aaggaaatagc

40261 tttgtat tttt tgtcattgat taaattgcac cagaaatggt tatggaaata ttaagaacat
40321 tttataaaac atacgaaaaa gtgattgtga aggattttta tttgcatgat tatagttaag
40381 caaatttcat ttcacatttt tgtaatgctg tacagcagtt cacaaaaaaa tgttcattgt
40441 agattttggt atgttcaatg ccaatgagtc tgtaatttta cgacctgtct gttctttttt
40501 tgggtggatt acgatgtaaa ttttctttct ctttctttct tttttttttt ttttgnntag
40561 aaaaaatagg gcaataatga tcaaagtttt gatgtcttga gtctccatct taggggatta
40621 tcttacgttt aagcttaaca tttcatgtag taagatttgg agagccacat actttacagt
40681 aaaattaagt gtgtataaaa cattaattgc tannnnnnnn nnnnnnnnnn nnnnnnnnnn
40741 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
40801 nnnnnnnnnn nnnngactaga tgttgctaaa aatctcacac tcggactctn anagaatacc
40861 cctattgtac aagtctgcaa aactagtatc ataacctcca ctagacatgt gaaaggtagt
40921 tcgtaatacg aactataaac aaatacacac acacacacac acacacacac acacacacac
40981 acacacagac tgaaagggtt tataagaaat tacagctaga aaataacaat atttaattga
41041 tgtat tttttg gccttcatat cttggtagtg gattgacatg aggtctacac tacaggccat
41101 agatggtttt aataaatcaa gagtctatac caatgatgag caatcttata aaactcctaa
41161 ttcaataaat tataggctga atttttttaa aataaaaagt tttattaaat cccatggaaa
41221 agatgaatag gcttaaaatc tgaattatca tctttacaga acagaggaaa gtccataaaa
41281 actagcactg ccttttgcca tcagggaac agctttcaaa cctttttttc tcctcaaatt
41341 agattatgtg cagctgtcca caaaatccca gctttctcat tattacccta tagtctcaca
41401 gcacataaga catcatacaa taaatgtaaa tacttgagtg actttaatat ctttccaata
41461 gaaatagaaa ttttaataacc tcaaataatta caatgtggtg attatattaa aatnnnnnnn
41521 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
41581 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnaactag gtgctgactc cggctcggtt
41641 gngagacact ttgccttatt ttagtttaat tatttaactt ttaattggct atgtaatgtt
41701 ttgaaaacct ggtttttaag aatgatgtaa ataagttcaa gagaggaaag ncaaaaacct
41761 aacagactct agttataaaa aaattgaaac attattcgct aataaaacca ccagtgtagc
41821 tctccctctc cctctcccta tccctctccc tctccccacg gncnccctct cctctctctt
41881 tctctccctt tctttttctc ttgtccctnn cttttctctt cttttccctt ctttntctt
41941 tcccccttcc cctccctttt ttccctnttc tccttttccc cctcccttcc cctctctct
42001 ccttcccccc cctctnccc ttccctttct ccnctttccc cccctcttcc tttctctttt
42061 tctccctttt cctctctctt ttccccctt cccctctctc cccccccctt ccccccttn
42121 ttcccccttt ttcttccctc ctctcttttt cctccccctt tctttttccc cctctttctt
42181 tccntttccc ctttccccct cccgcgcgc cgttcccttc gccccctttg ccgttgccgg
42241 tggcccgccg gggcccgccc ggctgccgct ctggcgccgc cccgcttttt tccnnnnnnn
42301 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
42361 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnggacc tacgggtctn gagnggcggg
42421 acttcagctc ngnggatgac actttccaca atactcaca atcagtaata taaataccat
42481 tctataaaac taggcaaata ttataaacac aacttacaaa agaaatatga atggccaata
42541 atcccaagga gaaaaaatg gtcagatgac tagtgactaa caaatataaa tactatcaaa
42601 ataatttgtt ttttctttgt ttttacttgg gntnngntn nttttttgtt nggtggtgtt
42661 gtgtttttnt gngnggttgt ggggtttttg tgtntttgtt ggggtgtttt gtgtnttggt
42721 ntgtgttgtg ggggttttgt ggggtggggg ttggtggng ggggtggtgt tgtggggggt
42781 tgggttgggg ttgggttttt gtgggtgtgg tgggtggtgt tgggtgtttgt tgggtgtttg
42841 ggggtgtgtt ttgggtgggt gtgttttttt tgtggtgttg gttgtttgtt ggggtnttg
42901 ggggtggggg gtgggtgggt gtgggtgtgt gtgggtgggt ggggggtttt tgttttgttg
42961 ttgtgtgggt ggttgtgggt tgtgttgggt gtttgggggt gtttttgggt gtttcgggtgt
43021 gttgtgggtg ggttgggggt tttttgtgtg ggtggggggg tttgtgggtt tgggtgggtgt
43081 ggcggttgtg cccnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
43141 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnngaacta
43201 cgtgctgaaa actagatttc tgtangnaa acccaattca cttgtgtctc ataatactca
43261 tctctctgag ggcacttcag tcaaaaactag cttattaata agattcttct agcacctcta
43321 aaactggacc taacgttaca cattatccat gccacaacaa atagcaaaag gcctccatta
43381 caatcatact aataggctct aagagctcaa aaattcaaaa cagttctatg aaaaatagag
43441 tagactttgt tctgtactat cccaaaagggt aaaagcagat gcaaaattta taggataaca
43501 gaataattaag aattaataaa aacctgtttt cccaactttt gggggatggg aacaccaagg
43561 tataagtgtg gtacactata ggtatgggta tgaagtacat agttgacttt tccaacaaat
43621 gaaatacatc atgagctctc tgcttaaagc attttcaacg tggtgaccag ctctcattaa
43681 taataaagga atttctgcag taagtgaaga ctaagaagac ataatacctt ttccaattca
43741 aggattttct tctgtaaatc cagataacaa gattgtttcc aagctaagtg atcaattctt
43801 aattttgctc atgtgaaaaa tttaaacagc caataaacat cgtaacttac agtatacata

43861	gttagaattc	aataaaat	tagaataatt	ctttcacagc	gggcagtatg	gctcnnnnnn
43921	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
43981	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnngaact	atgtgctgaa	acctgatttc
44041	tgtngnggnt	aacactctcc	ttggtttatc	aaacgtgcgc	acagtccatc	cagagggcct
44101	tacagtgtac	atatcaaccc	atatctcatt	cccttcttta	tcgggttaca	gaaccggttt
44161	acgcagtttc	ggcttagtga	aacaaaagaa	atcaccaatc	cgtatgccat	gcgtttatac
44221	gaatccctgt	gtcagtatcg	taagccggat	ggctcaggca	tcgtctctct	gaaaatcgac
44281	tggatcatag	agcgttacca	gctgcctcaa	agttaccagc	gtatgcctga	cttcgcgcgc
44341	cgcttcctgc	aggtctgtgt	taatgagatc	aacagcagaa	ctccaatgcg	cctctcatac
44401	attgagaaaa	agaaaggccg	ccagacgact	catatcgtat	tttccttccg	cgatatcact
44461	tccatgacga	caggatagtc	tgagggttat	ctgtcacaga	tttgaggggtg	gttcgtcaca
44521	tttgttctga	cctactgagg	gtaatttgtc	acagttttgc	tgtttctctc	agcctgcatg
44581	gattttctca	tactttttga	actgtaattt	ttaaggaagc	caaatttgag	ggcagtttgt
44641	cacagttgat	ttccttctct	ttcccttcgt	catgtgacct	gatatcgggg	ggttagttcg
44701	tcatcattga	tgagggttga	ttatcacagt	ttattactct	nnnnnnnnnn	nnnnnnnnnn
44761	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
44821	nnnnnnnnnn	nnnnnnnnnn	aactctatct	tgctgacacc	cagtcttctc	tanagnaccc
44881	ccatttgctt	tgactcctgt	tccttacacc	tgtatccagt	ctgtcatcaa	atcctgttga
44941	gtttacctcc	tgattatccc	ttgagttttt	ctactttttt	ccaccggctc	cccttaagtt
45001	ctcatcattt	ctttgcctct	ttgttggtct	cacattttact	cttgccccac	ctcaaattca
45061	ttttcataat	gtagccagat	tgtgaagtac	catgatttta	tctaccccc	agcagcagca
45121	gcagcaacag	tagcaaataa	tataaacctt	atatctcttc	aataatttac	actgcctttt
45181	aggataatgt	gccccacgaa	gctctgcata	atgggtcact	cttggttagc	taattcttcc
45241	ttcatatttt	gtatttttatt	tcattgattt	ctttttctaa	tttttatctt	ttgcctgtac
45301	caatatctaa	ttcactatat	tgataagtgt	ttatatttgg	tccaaaaatt	ctcaaagtct
45361	taatttttaa	aatgtcttgg	tcattccttg	gctcttcttg	ttattttcgc	ttattcagta
45421	ttgagtcctc	ccactgagat	agattttttt	aaggccttcc	tttcaataaa	gtttttattt
45481	tcttcataag	aataggcttt	gcacctttta	aaatgtttnt	atgggtttta	ttannnnnnn
45541	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
45601	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnaacacat	ttgctagcac	gcagtcctgc
45661	tgctganagn	atccccctct	ttacatactc	agttatttta	aaagaaaaca	actgaaacaa
45721	tgaacataaa	gaaaaatttt	aagaattctg	cagtactaag	tcaatgaaaa	taaataacca
45781	cctatgctga	ctgcttttcc	tattaaaaag	aatttcttat	tccctcatta	aattgacttt
45841	ctaccttcaa	gttcaagtgt	tttccttgat	aacaatcata	ccaatataga	aatgaaaaaa
45901	tacattcaaa	ctgaaaaata	catgaaaccc	taaagaaaaa	actcagcatc	tactaccct
45961	gaaacaatcc	tgtcaacatt	ttggtgaaca	ctcctaaaaa	ttctctagac	atattattat
46021	acagacatta	ataagtaagg	atataaaaata	cacagtattt	ggaaacattt	ttttcaacca
46081	acaatgtgtg	tagaacctca	tttgtcaaat	attgctaagc	actattcttt	ttaacagtc
46141	catagctctc	agtgtatgta	tagtgtatgg	ctataccaaa	actaatctca	tactgataag
46201	catttttagat	ttaaaaat	ttaaaaacac	cttaacaaca	ataccatgaa	tatccctgtg
46261	ctggctgttt	acatactctc	tctgagcttt	aaacctatcc	tttcataatc	tgctcagaga
46321	tctgaagct	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
46381	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
46441	tgacactcta	gtctttctct	nanagnaacc	cccaccgcgt	ctcccttaag	ttctcatcat
46501	ttctttgcct	ctttgtttgt	ctcacattta	ctcttgcccc	acctcaaatt	cattttcata
46561	atgtagccag	attgtgaagt	accatgattt	tatctacccc	ccagcagcag	cagcagcaac
46621	agtagcaaat	aatataaac	ttatatctct	tcaataattt	accatcgctt	ttaggataat
46681	gtgccccacg	aagctctgca	tcatgggtca	ctcttgttta	gctaattctt	ccttcataat
46741	ttgtatttta	tttcattgat	ttctttttct	aattttttatc	ttttgcctgt	accaatatct
46801	aattcactat	attgataagt	ttttatatatt	gggtccaaaa	ttctcaaagt	cttaatttta
46861	aaaatgtctt	ggtcattcct	tggtctctct	tgttattttc	gcttattcag	tattgagtct
46921	tcccactgag	atagattttt	ttaaggcctt	cctttcaata	aagtttttat	tttcttcata
46981	agaataggct	ttgcaccttt	taaaatgttt	tatgggtttt	attattattg	cttacattgc
47041	tatttccatt	attttcta	tggtatagct	gctatgcttt	tctttctttc	ttttctttnt
47101	tttttttttc	ttttgagaca	gtcttgcctc	gtgcagtnnn	nnnnnnnnnn	nnnnnnnnnn
47161	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
47221	nnnnnnnnnn	nnnnnnnta	cttcttgcta	gaaacccagt	ttgctctana	gaacccccaa
47281	ggtaacattt	tttggccagc	actttgaaat	ggccctctgt	gaagaacggg	tctctagcat
47341	acttataaac	gggcggtaac	ttttttactt	tggaatgacc	tgcaaggagc	tgcattcttc
47401	agaacaggtg	aaaaggcata	ggctctggaa	tcagaaaaag	cacgttttagc	tgcaggctct

47461	taacacaggg	ggccatattc	atctctgtat	tgttccagtt	ttagttcatg	tgctgccaa
47521	gcaagcacag	ccatagcctt	ttaacatggt	tcctggcttt	actaataaat	ccacaaaaat
47581	gggaaagatt	taagacttaa	tatttttaaa	atgcagatct	gattctgtca	ctctcaccce
47641	taccctgatt	tcctcagtgg	cttcccactg	cacttggaat	aaattcacac	tctttgccat
47701	ggcctaacgg	ccctgcacaa	cctggtaact	tctactctca	gacctcactt	tattcctcct
47761	ttccctggcc	cccagctcca	gttatagtct	ctcttcccta	gaacacccca	gacttggtct
47821	gcctctgggc	tattgcacct	gtgattccct	ttgctcgga	ctctctttcc	ctagagcttc
47881	agctcttcat	atgcctgcct	gtcacccttc	agatatcagc	tcagctgnnn	nnnnnnnnnn
47941	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
48001	nnnnnnnnnn	nnnnnnnnnn	nnnnnnngac	tagttgcaga	actctagttg	acttaagaac
48061	cctaacatat	tttgactggg	cactcacact	ttctacgaga	tttctaaatg	taagaatcta
48121	ttgtttaaaa	tattttacaat	gtataacagt	accctgcacc	agacatgtga	ccgccactgt
48181	aaaataacca	ctgtacaaaa	gcataattcca	gtaaaaatat	atgcacatgt	acaaagcaat
48241	tctaatttaa	aaaaaaaaaga	aaaagaaaca	gcaccattga	acagttacac	caagagctga
48301	aattattatgc	aatccaaaga	actgtgaaat	agtaattgca	tacctaat	tcccctgttc
48361	tgtcaagtct	ccatcactat	gcaatattgt	tgtagcaat	cttaaggtag	aagttcctga
48421	tttactgtga	ttatttttca	ttccaagtag	ccatcgaacc	atcattttaa	tagcctgaat
48481	ctattttaaga	gacaagtcga	caaataaaac	actggtacat	tttgaaagaa	tttctgtatc
48541	aacttttttt	aaaaataaaa	tgtatttaaat	ggatgctttt	taagaaaaga	aaaaattctt
48601	ttctgtatat	atgaagataa	tctcgaatgt	tttaaagatg	tttcaaacat	ccaagaaaaat
48661	ggcaaaagaa	tgatttgact	aaagactagt	ttgttctttg	ctgttttaan	nnnnnnnnnn
48721	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
48781	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnna	ctaactgtg	ncaaaaaacc	ttacactttg
48841	gtngnagaat	aacacttata	acttgtatta	taacatgttt	ataggaacta	tcctttcaag
48901	tttttaataa	aattgtatta	accagcaagg	caatctgtgc	ctgtttattg	cctttataaa
48961	taattctcta	aaaaagggaa	aaaagttata	ctaggaaaag	cagtattaca	ataggaaaaa
49021	tctctattta	atattagtct	gaattgcagt	tctgttgcta	agcaaagtga	agaactccag
49081	ttacagcctg	gatattctta	tatttgttac	ctgacatttt	aaatagacaa	ttagttccca
49141	caaactctct	tgttaccaa	ccacaagttt	gcagcaactc	aaagaaaaag	acagctacca
49201	tcaggtttta	gtcagtggcc	tgtccaaatg	ttgaagagga	cctgtcctct	aaccttttagc
49261	tcccaactga	caagtcaccc	atcggactct	atctcaaaaa	aaaaaaaaaa	aaaaaaaaaa
49321	aggctatttg	gacgcaaact	tgtttaataa	aagagaagat	cacaacaaag	accccgatac
49381	catcttctac	tactgtcgct	aataggagtt	ctggggaact	tgtgagaggt	ggaagtacac
49441	actgttctct	ttttcaataa	tcaagatatt	tagtttactt	aggagcttat	tcaaaaactg
49501	attcacatan	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
49561	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnng	aactacgtgn
49621	cagaaaaccn	agactgacgg	tnanagnaaa	caccggcagt	atagtgtctg	aaagctttta
49681	cacattagta	taattttatt	ggatgtagat	aaaccaatag	taaagataaa	atcataatac
49741	ttgatttaac	tcaaggaaaag	cagacataaa	gaatagaaga	ggaaaactat	caagatggta
49801	gctttaaac	ccaccatata	aataattaca	tgaagtgtaa	atgctctaag	cactcccatt
49861	gaaacagttt	gtcaaattgg	attcaaaaaca	accaatataa	aagcaaacaa	ggaaacaaaa
49921	aacaacaaca	gaaagacaac	taggacattt	atttaaaaaat	aaataaataa	ataaactttt
49981	tctttaaggg	tggtgtctta	ttgccagggc	tggaatgcag	tggtacaatc	atagctcact
50041	ggagcctgaa	actcctgggc	tcaagtgate	ctccgcctc	atattctcaa	agcacggaga
50101	tttacaagtg	tgagccacca	tggccagcca	agaaactcac	tttaaataatg	aagtcccagg
50161	aagactaaaa	atggaaaagga	aaaatatatg	ccttacaac	actaggacca	aagaaaagca
50221	caaattggcta	cactattctc	agacacaaaa	gcctgcagaa	caaggaagat	gaccagagtt
50281	aggaagagac	atatttcata	acaacatcgt	ctactggcaa	ggctgtgggg	annnnnnnnn
50341	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
50401	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nngacttacg	ttgcagaaat	ctncatttga
50461	ctctnanaga	ataccctttt	cccatccac	tcacagtctg	ttagcaccat	tggatatggc
50521	accaccgccc	tggggctgtg	cagcttgacc	ctgccttctc	catcctgccc	caactgtagc
50581	aaacgtttca	caagcgctc	tgccggggaa	ggagagggcg	aaagctcaga	agcacagtct
50641	gctgtcttct	ccccggggcg	acagtccacg	tctcgagagc	gcgcccggctc	ccagggtctct
50701	gctccccctgt	ctcactgtgc	acggcatatg	aaaactgaag	aggacaacgg	gctgtttctc
50761	aatactgcat	ttcctaagag	atccaaaggg	aatttcttct	attttcgttc	ttaataattta
50821	aaaaaaggac	cttagagttc	tggttttgta	gccagagaa	acctatggat	ttaatcatca
50881	agaattatca	gtagacagaa	aggggtagtt	gtggacctta	aggtttgacc	agaaaacacc
50941	ccgtctgagc	ccccactacc	ccactcacc	cccaacacac	acatatacac	tcatacacac
51001	acatacacac	acacatatat	acgtgcacac	acatagacac	acacacgcac	acacacatat

51061	acacacacac	atatacacgc	acacacttac	ccactcacac	acatacaggc	acatgcatac
51121	acacacaagt	acaccnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
51181	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnngacc
51241	tacggctcntg	aggtgtcgat	tcagtctgna	gntaccacct	caacctatgct	gatacaaagc
51301	actcttcctg	tggctaaaagc	agggatgcaa	tagagaggca	gctgattgca	ggtcagagac
51361	tagaaaggag	cagcagggtcc	caggaatggg	tgttgttttg	tttgtaaagc	cattcccagg
51421	ttatgtgaca	ctagcaacag	ccctgggtgg	tttttttatt	gaacacacat	cgagatttat
51481	ttgaggccca	tacaattgac	attcaaacc	tcaccttgaa	gaactcaact	cattatactt
51541	cgatcacggc	tactggagga	ctacccgagg	gtgagtttgt	tgatcgcccc	cttcggcctc
51601	ctcggcgccg	cccaagttcg	acattcagct	tgatccccc	accagagccc	ggttcgggtc
51661	cctttccccc	accgggttta	ctattcgc	ttcccttggt	ttccctgtc	ccctgtccct
51721	atcggtgttt	ttaccacgtc	tcccactccc	tttttcccc	tgatccctcc	gcccccccc
51781	cgtcccttcc	cccttttctc	ccctttttcc	tttccggttc	acgtttttcg	cccccatat
51841	tgatattgcc	ccttttctct	ccccctcccc	ccttgcccat	cctccttccc	ccccatgcct
51901	tccctcccg	ttttcccttt	ctttattccg	ccccattcct	cacctttcca	ccccattttt
51961	ccccccccc	catttttccc	cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
52021	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
52081	ngactgagtt	gcagagtctn	agctcngctc	tnanagaata	cccctcatct	ctcttggggt
52141	ctggcatcac	attatgttct	ctttatgtgt	aaatatattt	aggagcaaat	caaataccac
52201	tttcaatttt	caaataagct	tactcggttt	ataattaaat	ttctcagagc	agtttgtcaa
52261	ccttggcatt	tacacataaa	acactaaca	ttgtgtaaag	attcttaaga	gatttgttta
52321	gaatatgttg	agttaagaag	cacaatacca	gagaaaaggga	aaattaaaa	actgaagtag
52381	tcgtattttac	atcacagctc	tgagaacaaa	agcagtacag	cagagcacac	cccagcagaa
52441	gtgattcatg	ctcgtcagg	ccgaggcaga	gtgaggggagc	cataggggag	agagagcatg
52501	gaaaccaagg	ccagcacaat	attcctacta	gatggatgga	gggagatggg	aagagaggga
52561	agcagagcta	gatacagaga	gagagagaga	aagagagaga	gagagagaca	gagaaggaga
52621	cagagttaag	actaatgctg	gcactataat	ccagtaaaca	aaaatacatc	cttattctct
52681	cacacagaaa	tataacctgg	gttggtggatc	tccgtgaggc	cagaccattg	actgcgccgc
52741	aggcccttgg	gggggaggcg	aggggcggag	ccgtgggttc	gcgnnnnnnn	nnnnnnnnnn
52801	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
52861	nnnnnnnnnn	nnnnnnnnnn	nnngaactag	tgctgaatct	acctcggtgt	anagaaaacc
52921	ttgccctgca	ctccagcctg	ggtaacagag	tgagacactg	tctcaaaaa	aaaaaaaaaa
52981	aaaaaaaaaga	agtggtttat	acttggcaga	aaggatttca	aatagggagc	ttccttcaca
53041	gatgtatgtg	tttaaaggct	agtcaatgag	tagtagagct	gttttgagct	caagaacatg
53101	gagccacagc	taggagaaga	ttaaattgaaa	aaaattccca	aaaaagaaaa	gaaaaaggaa
53161	aatacccccc	accaagatct	cagtcacttt	ccaaatcaag	atccccca	catcttgctc
53221	agcttctcat	cctcaactta	gatgggtgcca	gagatcccaa	tcaagatcat	atctagaagg
53281	cagaaatctc	caggaagaac	tccaccaaga	cgaatgcctg	ttctaccag	gcctaggaga
53341	aatggctgtg	cagtgggagg	atgaacgc	tcatacagcat	tcttgctccag	gagttagctca
53401	tcatacctctc	ctgtcatttc	tggtcaccac	caaagaagtc	ttgcaaaagg	acatctagtc
53461	tctagttaag	gcaattaact	agtgaattaa	ccattgaatt	gcacactaaa	cgggtgaatt
53521	tgatgggata	tgaattatac	ctcaataaag	ctgcatatat	actgcgcctg	tatgtgcggc
53581	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
53641	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn
53701	tgcggttagga	gtaagctaac	ttcttaaaat	tttaaattgta	aacatactta	ttatttgtaa
53761	ctcctaaaca	ttggcatata	taaattaggg	gaacagagac	atgaaaacag	aatagattaa
53821	ttagaataaaa	agaaagtaaa	cagatacaga	atgaatagct	taccaagt	ttttttatat
53881	gttttggttg	agatggatgt	gcttctgtta	acattctatg	caaagaatac	agtataagca
53941	aaacatacaa	ggagaataca	gtataagcaa	aacatacaag	atggaaaggc	agaaggcagc
54001	tggcaggcat	ctttgtatgt	tttcttatta	aagtgtgtgg	ctctgctgag	gggaagcggg
54061	actgtggccc	ctggggtgtc	gcgcgtacct	tcccgctccg	gcccccggtt	ctctctccgt
54121	tggcggtttt	cagcggcccg	aggggtgtcta	gtttgctgcg	tggcacggtc	tcccggttcc
54181	tgggttgcgtg	tcgtcccggc	cgcgtcttct	cctcattccc	cgtgttttgt	tacgctagtc
54241	ggttggcccg	ccccccgggtg	tttgttgccg	ttttgggttg	cctgtgtcgc	gttcccgcgc
54301	cccggccccc	gtatccgtgc	cccccgcttc	tcctattcgc	tttgacctgc	gatccccccc
54361	ttttcgtcgc	ctttttttgc	cgggcccggt	tgggcccc		

=>

LOCUS (LOC): AB014548 GenBank (R)
 GenBank ACC. NO. (GBN): **AB014548**
 CAS REGISTRY NO. (RN): 211153-78-5
 SEQUENCE LENGTH (SQL): 5177
 MOLECULE TYPE (CI): mRNA; linear
 DIVISION CODE (CI): Primates
 DATE (DATE): 6 Feb 1999
 DEFINITION (DEF): Homo sapiens mRNA for KIAA0648 protein, partial cds.
 SOURCE: Homo sapiens adult male brain cDNA to mRNA,
 clone_lib:pBluescriptII SK plus clone:HJ03994.
 ORGANISM (ORGN): Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Euteleostomi; Mammalia; Eutheria; Primates;
 Catarrhini;
 Hominidae; Homo
 NUCLEIC ACID COUNT (NA): 1690 a 931 c 1037 g 1519 t
 REFERENCE: 1 (bases 1 to 5177)
 AUTHOR (AU): Ohara,O.; Suyama,M.; Nagase,T.; Ishikawa,K.
 TITLE (TI): Direct Submission
 JOURNAL (SO): Submitted (26-MAY-1998) to the DDBJ/EMBL/GenBank
 databases. Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology; Yana 1532-3, Kisarazu,
 Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
 Tel:+81-438-52-3913, Fax:+81-438-52-3914)
 OTHER SOURCE (OS): CA 129:171313
 REFERENCE: 2 (sites)
 AUTHOR (AU): Ishikawa,K.; Nagase,T.; Suyama,M.; Miyajima,N.;
 Tanaka,A.; Kotani,H.; Nomura,N.; Ohara,O.
 TITLE (TI): Prediction of the coding sequences of unidentified
 human genes. X. The complete sequences of 100 new cDNA
 clones from brain which can code for large proteins in
 vitro
 JOURNAL (SO): DNA Res., 5 (3), 169-176 (1998)

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..5177	/organism="Homo sapiens" /db-xref="taxon:9606" /clone="HJ03994" /clone-lib="pBluescriptII SK plus" /dev-stage="adult" /sex="male" /tissue-type="brain"
gene	1..2558	/gene="KIAA0648"
CDS	<1..2558	/gene="KIAA0648" /codon-start=3 /product="KIAA0648 protein" /protein-id="BAA31623.1" /db-xref="GI:3327110" /translation="LYYLYASLDPNAVKALENMW KCQNMLRSHVRELLDLHKQPTSEA NCSAMFGKLMNTIAKNLPDPGKAQDFVKKFNQVLG DDEKLRSQLELLISPTCSCKQADI CVREIARKLANPKQPTNPFLEMVKKFLLERIAPVH IDSEAIKLVKLMNKSIEGTADDE EEGVSPDTAIRSGLELLKVLSTFTHPTSFHSAETY

ESLLQCLRMEDDKVAEAAIQIFRN
 TGHKIETDLPQIRSTLIPILHQAKRGTPHQAKQ
 AVHCIHAIFTNKEVQLAQIFEPLS
 RSLNADVPEQLITPLVSLGHSMLAPDQFASPMK
 SVVANFIVKDLLMNDNSTGEKNGK
 LWSPDEEVSPEVLAKVQAIKLLVRWLLGMKNNQS
 KSANSTLRLLSAMLVSEGLTEQK
 RISKSMSRLRLAAGSAIMKLAQEPYHEIITPE
 QFQLCALVINDECYQVRQIFAQKL
 HKALVKLLLPLEYMAIFALCAKDPVKERRAHARQ
 CLLKNISIRREYIKQNPIMATEKLL
 SLLPEYVVPYMIHLLAHDPDFTRSQDQDLRDIK
 ECLWFMLEVLMTKNENNSHAFMKK
 MAENIKLTRDAQSPDESKTNEKLYTVCDVALCVI
 NSKSALCNADSPKDPVLPKFFFTQ
 PEKDFCNDKSYISEETRVLLLTGKPKPAGVLGAV
 NKPLSATGRKPYVRSTGTETGSNI
 NVNSELNPSTGNRSREQSSEAAETGVSENEENPV
 RIISVTPVKNIDPVKNKEINSDQA
 TQGNISSDRGKKRTVTAAGAENIQKTDKVDDES
 GPPAPSKPRRGRPKSESQGNATK
 NDDLKPKPINKGRKRAAVGQESPGGLEAGNAKAPK
 LQDLAKKAAPAERQIDLQR"

SEQUENCE (SEQ):

1	gcttatatta	cttatatgct	agtttggatc	caaatgctgt	aaaagctctc	aacgaaatgt
61	ggaagtgtca	gaacatgctt	cggagccatg	tacgcgaact	attggatttg	cacaagcagc
121	ctacatcaga	ggctaactgt	tctgccatgt	ttggaaaact	gatgaccata	gcaaagaatt
181	tgcctgaccc	cgggaaagca	caagattttg	tgaagaaatt	taaccaggtt	ctcggcgatg
241	atgagaaaact	tcggtctcag	ttggagttat	taattagccc	aacctgttct	tgcaacaag
301	cagatatattg	tgtgagagaa	atagcccgga	aacttgcaaa	tcctaagcaa	ccaacaaatc
361	cttttctaga	gatggtcaaa	tttctgttgg	aaagaatcgc	acctgtgcac	attgattcag
421	aagccataag	tgcactagt	aaattgatga	ataagtcaat	agaggggaca	gcagatgatg
481	aagaggaggg	tgtaatgtca	gatacagcta	tccgttcagg	acttgaactt	cttaagggtc
541	tgtcttttac	acatcctacc	tcgttccact	ctgcagagac	atatgagtcc	ttgttacagt
601	gcctaagaat	ggaggatgac	aaggtagcag	aagctgctat	tcaaattttt	agaaatacag
661	gtcacaaaat	agaaacagac	cttccccaga	tacgatcgac	cttaattccc	attttacatc
721	aaaaagcaaa	gaggggtact	ccacaccaag	caaaacaggc	tgtgactgt	atacacgcca
781	tattcacaaa	taaagaagtc	cagcttgcac	agatttttga	gccactcagt	aggagtctga
841	atgctgatgt	gccagaacaa	cttataactc	cattagtttc	attgggccac	atttctatgt
901	tagcaccaga	tcagtttgct	tccccaatga	aatctgtagt	agcaaatttt	attgtgaaag
961	atctgcta	gaatgacagg	tcaacaggtg	aaaagaatgg	aaaactgtgg	tctccagatg
1021	aagagggttc	ccctgaagta	ctagcaaagg	tacaggcaat	taaacttctg	gtaagggtggc
1081	tgttgggtat	gaaaaacaac	cagtctaaat	ctgccaatc	aacccttcgg	ttattatcag
1141	cgatgttggt	tagtgagggt	gacctgacag	agcaaaaagag	gatcagtaaa	tctgatatgt
1201	ctcgcttgcg	attagctgct	ggtagtgcc	taatgaagct	tgctcaggaa	ccttggtacc
1261	atgaaattat	tacccagaa	cagtttcagc	tctgtgcact	tggtattaat	gatgagtgtt
1321	accaagtaag	gcagatattt	gctcagaagc	tgcataaggc	acttgtgaag	ttactgctcc
1381	cattggagta	tatggcgatc	tttgccctgt	gtgccaaaaga	tcctgtgaag	gagagaagag
1441	cacacgcacg	acaatgttta	ctgaaaaata	tcagtatacg	cagggaatac	attaagcaga
1501	atcctatggc	tactgagaaa	ttattatcac	tggtgcctga	atatgtagt	ccatacatga
1561	ttcacctgct	agcccatgat	ccagatttta	caagatcaca	agatgttgat	cagcttcgtg
1621	atatcaaaga	gtgcctatgg	ttcatgcttg	aagttttaat	gacaaaagaat	gaaaacaata
1681	gccatgcctt	tatgaagaag	atggcagaga	acatcaagtt	aaccagagat	gcccagctctc
1741	cagatgaatc	caagacaaat	gaaaaactgt	atacagtatg	tgatgtggct	ctctgtgtta
1801	taaatagtaa	aagtgccttg	tgcaatgcag	attcaccaaa	ggacccagtc	ctccaatga
1861	aattttttac	acaacctgaa	aaggacttct	gtaacgataa	gagttatatt	tcagaagaga
1921	caagagtact	tctgttaaca	ggaaagccaa	agcctgctgg	agtactaggt	gcagtaaata
1981	agcctttatc	agcaacggga	aggaaacctt	atgttagaag	cactggcact	gagactggaa
2041	gcaatattaa	tgtaaattca	gagctgaacc	cttcaaccgg	aaatcgatca	agggaacaga

2101	gttcagaggc	agcagaaaact	ggagtttagtg	aaaatgaaga	gaaccctgtg	aggattattt
2161	cagtcacacc	tgtaaagaat	attgacccag	taaagaataa	ggaaattaat	tctgatcagg
2221	ctacccaggg	caacatcagc	agtgaccgag	gaaagaaaag	aacagtaaca	gcagctgggtg
2281	cagagaatat	ccaacaaaaa	acagatgaga	aagtagatga	atcgggacct	cccgcccctt
2341	ccaaacccag	gagaggacgt	cgaccceaagt	ctgaatctca	gggcaatgct	acccaaaaatg
2401	atgatctaaa	taaacctatt	aacaagggaa	ggaagagagc	tgcagtgggt	caggagagcc
2461	ctggggggtt	ggaagcaggt	aatgccaaag	cacccaaact	gcaagattta	gccaaaaagg
2521	cagcaccagc	agaaagacaa	attgacttac	aaaggtaaaa	atgcatttgc	aaaggggagaa
2581	aatgaaggcc	aaacagaagc	aggctccagc	ttctgcaaaa	acttggattc	acaaatgtcc
2641	ctgaacagaa	aatgaagctc	acttcagaac	acacactctc	tgccttgaaa	actaaagaga
2701	ctattacttc	cttttcacat	gaccacaagt	cctctgatgg	aaatgtacag	cagaaactct
2761	tgagagagag	gctaaaaagca	actctgttct	cccccttccc	ctagactttt	cttacgaaaa
2821	gtcaataatt	aagcaaatgt	cttaacactt	ggttccagtt	cctgcctatc	tggagtttaa
2881	atgctgaata	caccattaat	ttccacgctg	cagttttttat	tttaaagaaa	gtaacaagat
2941	gtcttttacac	tgacactgaa	aattcatcca	ttttagagcc	aggaattccc	atgttacaca
3001	ggaaaaaata	gaagtctact	gaattaattt	tttaaaagaa	aagagatcag	attaaatatt
3061	tctttgtttt	tccttttgga	aacttttatg	tataattctt	tctgcctgcc	tacttttctg
3121	caaaaatgag	atgtacagat	ttcggttccc	tgctatgaaa	agtgatgtgg	tagcaatttt
3181	ataaatgttg	ctttctgatt	tttatcagag	tgagaaaatt	aaaattattg	atttgcaagt
3241	agtaaacagt	tcataatttg	atttccccct	attttagttt	aatataattt	gcaataaatg
3301	tacatattgt	tgtttgtttc	ataaaagcata	tcactttaaa	atgggtttta	ctcctgtgat
3361	tatgtttgga	tatttggaat	tttaaaggag	taaagactgt	ccagcatttg	gttttataat
3421	gtttgtcacc	agattttttat	taatgtaaaa	aatcaatttt	taaaaaatag	ttggactttg
3481	gcagctttta	aggaaagtgt	gaggtgtttt	aggattgcta	tcaattttca	gcattgtgct
3541	atttggaagt	aagtgttttg	cttttgtctg	atgggtctgg	ctcattttta	tgtttatatt
3601	agaaaactgt	tgcatcaata	tattatgttt	cttggcattg	ttcagcatag	gtaatgtgtg
3661	cactttatgt	gtacacataa	tcataattta	gttttttgca	taaaaataat	gcttctagat
3721	gtcatggcag	tctttttaat	ctttttatca	tatgctttct	tgtgaatttt	ttcatgttaa
3781	agagctaaag	tcataacatg	attacagtca	actctccatt	atctatataa	aatagtgact
3841	aagcctcagg	tttttaattt	tgtgataaca	aaataacgaa	ggcatgtaac	acctgattct
3901	ggaggaacat	gaaatttgtc	ttttctcatg	tccagagttc	tatcctgccc	ccactgtcca
3961	ctgtagggtc	atccgcaaag	ccctagcaga	atgtgctcac	tccatttccct	tacacgtttc
4021	tagcatgggt	cagaggaaac	aacatttgtg	ttataacttc	gtcttgatag	gctgtagtgt
4081	acatgggatg	taaaacaaac	aagtgtatca	aagggtggatg	attctgttag	agtgaagtgt
4141	gagagttaa	gtcacttacg	tttctcatag	atgatcaaga	gttggctgtg	tattgactga
4201	aagatgggta	attattttta	atatgcattt	acacacattt	aggtatcaga	agatgcttag
4261	ggaacaatgg	ataccaatga	tagaaaaatga	tacctttaca	ggggcagaaa	aatccccact
4321	cttccttatt	gcctcttcag	aaccctttag	aaagtataaa	atattgcctc	caacatgctg
4381	aaaaagagta	tctatgcata	agtatcagag	aagtccctca	agcaatcagt	aggtgtgttc
4441	tatttagaga	gagtttaaag	ttctcttagc	atcagacaac	ttgattccta	aggtttccag
4501	tgtgtcacca	acaaaaagtg	cattgatagg	gacctttgtc	tcttcctccc	tttgattaat
4561	tgcccgcat	cacagtttac	tagattacca	agtgttacat	catattaaat	aaaatgtagc
4621	agaaccatct	gcatcaatat	attcctgttt	agatttttgc	aggagagaag	ttaaaaggat
4681	ttgctccttg	tatgatgtaa	gtggcccacc	ccaattttgt	aacatgatgc	aagtgtctgg
4741	cactaaggga	agcaagagta	gggttgtgga	aagaccaagc	tgatggggag	ggacttgttt
4801	acgggaattt	tttttagttt	ccttttcaaa	ggaaaacatt	aaaaatccct	aggaatttgg
4861	tattcacatc	tcagagaact	acaacacaaa	agtgcagact	tatatattgag	aattaatgtt
4921	aaccctttgt	gtctagtttg	aagcttcttg	tatttgtcta	aaacaacaag	ccagaatttt
4981	gtatctcctt	tgataaaaaag	tgtgtataat	gtaaagtagt	tttgcatatt	cttgtgctgc
5041	acatgggctg	aattttttaa	ttttttttaa	aaacttgaag	cagaaccttg	taatttgtgt
5101	aatgacaag	tgtaaaatcc	taccataaaa	tgctaaaaat	atgcactggt	tcaaataaaa
5161	ccaagaaatg	cagcatt				

=>

LOCUS (LOC): AB023196 GenBank (R)
 GenBank ACC. NO. (GBN): **AB023196**
 CAS REGISTRY NO. (RN): 222526-42-3
 SEQUENCE LENGTH (SQL): 5110
 MOLECULE TYPE (CI): mRNA; linear
 DIVISION CODE (CI): Primates
 DATE (DATE): 16 Jun 1999
 DEFINITION (DEF): Homo sapiens mRNA for KIAA0979 protein, partial cds.
 SOURCE: Homo sapiens adult male brain cDNA to mRNA,
 clone_lib:pBluescriptII SK plus clone:hj07056.
 ORGANISM (ORGN): Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Euteleostomi; Mammalia; Eutheria; Primates;
 Catarrhini;
 Hominidae; Homo
 NUCLEIC ACID COUNT (NA): 1714 a 919 c 1045 g 1432 t
 REFERENCE: 1 (sites)
 AUTHOR (AU): Nagase,T.; Ishikawa,K.; Suyama,M.; Kikuno,R.;
 Hirosawa,M.; Miyajima,N.; Tanaka,A.; Kotani,H.;
 Nomura,N.; Ohara,O.
 TITLE (TI): Prediction of the coding sequences of unidentified
 human genes. XIII. The complete sequences of 100 new
 cDNA clones from brain which code for large proteins
 in
 vitro
 JOURNAL (SO): DNA Res., 6 (1), 63-70 (1999)
 OTHER SOURCE (OS): CA 131:28520
 REFERENCE: 2 (bases 1 to 5110)
 AUTHOR (AU): Ohara,O.; Nagase,T.; Kikuno,R.
 TITLE (TI): Direct Submission
 JOURNAL (SO): Submitted (04-FEB-1999) to the DDBJ/EMBL/GenBank
 databases. Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology; Yana 1532-3, Kisarazu,
 Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
 Tel:+81-438-52-3913, Fax:+81-438-52-3914)

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..5110	/organism="Homo sapiens" /db-xref="taxon:9606" /clone="hj07056" /clone-lib="pBluescriptII SK plus" /dev-stage="adult" /sex="male" /tissue-type="brain"
gene	1..4253	/gene="KIAA0979"
CDS	<1..4253	/gene="KIAA0979" /codon-start=3 /product="KIAA0979 protein" /protein-id="BAA76823.1" /db-xref="GI:4589602" /translation="VRRLKMOVVKTfMDMDQDSEE EKELYLNALHLASDFFLKHPDKD VRLLVACCLADIFRIYAPEAPYTSPDKLKDIFMF ITRQLKGLEDTKSPQFNRYFYLL NIAWVKSYNICFELEDSNEIFTQLYRTLFSVINN

GHNQKVHMHMVDLMSSIIICGDTV
 SQELLDTVLVNLVPAHKNLNKQAYDLAKALLKRT
 AQAIEPYITNFFNQVLMGKTSIS
 DLSEHVFDLILELYNIDSHLLSVLPQLEFKLKS
 NDNEERLQVVKLLAKMFGAKDSEL
 ASQNKPLWQCYLGRFNDIHVPIRLECVKFASHCL
 MNHPDLAKDLTEYLKVRSHDPEEA
 IRHDVIVSIVTAAKKDILLVNDHLLNFVRERTLD
 KRWRVRKEAMMGLAQIYKKYALQS
 AAGKDAKQIAWIKDKLLHIYYQNSIDDRLLVER
 IFAQYMPHNLETTERMKCLYYLY
 ATLDLNAVKALNEMWKCQNLLRHQVKDLLDLIKQ
 PKTDASVKAI FSKVMVITRNLPDP
 GKAQDFMKKFTQVLEDEKIRKQLEVLVSPTCSC
 KQAEQCVREITKKLGNPKQPTNPF
 LEMIKFLLERIAPVHIDTESISALIKQVNKSIDG
 TADDEDEGVPTDQAIRAGLELLKV
 LSFTHPISFHS AETFESLLACLKMDDEKVAEAAAL
 QIFKNTGSKIEEDFPHIRSALLPV
 LHHKSKKGPPRQAKYAIHCIHAIFSSKETQFAQI
 FEPLHKS LDPSNLEHLITPLVTIG
 HIALAPDQFAAPLKSLVATFIVKDLLMNDRLPG
 KKTTKLWVPDEEVSPETMVKIQAI
 KMMVRWLLGMKNNHSGSTLRLLTILHSDGD
 LTEQGGKISKPDMSRLRLAAGSAIV
 KLAQEPYHEIITLEQYQLCALAINDECYQVRQV
 FAQKLHKGLSRLRPLEYMAICAL
 CAKDPVKERRAHARQCLVKNINVRREYLKQHAHV
 SEKLLSLLPEYVVPYTIHLLAHP
 DYVKVQDIEQLKDVKECLWFVLEILMAKNENNSH
 AFIRKMVENIKQTKDAQGPDDAKM
 NEKLYTVCDVAMNIIMSKSTTYSLESPKDPVLPA
 RFFTQPDKNFSNTKNYLPPPEMKSF
 FTPGKPKTTNVLGAVNKPLSSAGKQSQTKSSRME
 TVSNASSSSNPSSPGRIKGRLDSS
 EMDHSENEYTMSSPLPGKSKDKRDDS DLVRSEL
 EKPRGRKKT PVTEQEEKLGMDDLT
 KLVQEQKPKGSQRSRKRGH TASESDEQQWPEEKR
 LKEDI LENEDEQNSPPKKGKRGRP
 PKPLGGGTPKEEPTMKTSKKGSKKSGPPAPEEE
 EEEERQSGNTEQKSKSKQHRVSRR
 AQQRAESPESAIESTQSTPQKGRGRPSKTPSPS
 QPKKNVRVGRSKQAATKENDSSEE
 VDVFGSSPVDDI PQEETEEEEVSTVNVRRRS AK
 RERR "

SEQUENCE (SEQ):

1	tggtgagacg	attaaagatg	gttgtgaaaa	cttttatgga	tatggaccag	gactctgaag
61	aagaaaagga	gctttattta	aacctagctt	tacatcttgc	ttcagatttt	tttctcaagc
121	atcctgataa	agatgttcgc	ttactggtag	cctgctgcct	tgctgatatt	ttcaggattt
181	atgctcctga	agctccttac	acatcccctg	ataaactaaa	ggatatattt	atgtttataa
241	caagacagtt	gaaggggcta	gaggatacaa	agagcccaca	gttcaatagg	tatttttatt
301	tacttgagaa	cattgcttgg	gtcaagtcac	ataacatatg	ctttgagtta	gaagatagca
361	atgaaathtt	cacccagcta	tacagaacct	tattttcagt	tataaacaat	ggccacaatc
421	agaaagtcca	tatgcacatg	gtagacctta	tgagctctat	tattttgtgaa	ggtgatacag
481	tgtctcagga	gcttttggat	acggttttag	taaatctggt	acctgctcat	aagaatttaa
541	acaagcaagc	atatgathtt	gcaaaggctt	tactgaagag	gacagctcaa	gctattgagc
601	catatattac	caattttttt	aatcagggttc	tgatgcttgg	gaaaacatct	atcagcgatt
661	tgtcagagca	tgtctttgac	ttaatttttg	agctctacaa	tattgatagt	catttgctgc
721	tctctgtttt	accccgactt	gaattttaaat	taaagagcaa	tgataatgag	gagcgcctac

781	aagttgttaa	actactggca	aaaatgtttg	gggcaaagga	ttcagaattg	gcttctcaaa
841	acaagccact	ttggcagtg	tacttgggca	ggtttaatga	tatccatgta	ccaatccgcc
901	tggaatgtgt	gaaatttgct	agccattgtc	tcatgaacca	tcctgattta	gcaaaagact
961	taacagagta	tcttaaagtg	aggtcacatg	accctgagga	agctattaga	catgatgtta
1021	ttgtgtcaat	agttacagct	gctaaaaaag	atattcttct	ggccaatgat	cacttactta
1081	atthttgtgag	agagagaaca	ttagacaaac	gatggagagt	acgcaaagaa	gccatgatgg
1141	gacttgccca	aatttataag	aaatatgctt	tacagtcagc	agctggaaaa	gatgctgcaa
1201	aacagatagc	atggatcaaa	gacaaattgc	tacatatata	ttatcaaaat	agtattgatg
1261	atcgactact	tggtgaacgg	atctttgctc	aatacatggg	tcctcacaat	ttagaaacta
1321	cagaacggat	gaaatgctta	tattacttgt	atgccacact	ggattttaat	gctgtgaaag
1381	cattgaatga	aatgtggaag	tgtcaaaatc	tgctccgaca	tcaagtaaag	gatttgcttg
1441	acttgattaa	gcaacccaaa	acagatgcc	gtgtcaaggc	catattttca	aaagtgatgg
1501	ttattacaag	aaatttacct	gatcctggta	aggctcagga	tttcatgaag	aaattcacac
1561	aggtgttaga	agatgatgag	aaaataagaa	agcagttaga	agtacttgtt	agtccaacat
1621	gctcctgcaa	gcaggctgaa	ggttgtgtgc	gtgaaataac	taagaagttg	ggcaacccca
1681	aacagcctac	aaatcctttc	ctggaaatga	tcaagtttct	cttgagagag	atagcacctg
1741	tgacataga	taccgaatct	atcagtgctc	ttattaaaca	agtgaacaaa	tcaatagatg
1801	gaacagcaga	tgatgaagat	gagggtgttc	caactgatca	agccatcaga	gcaggctctg
1861	aactgcttaa	ggtactctca	tttacacatc	ccatctcatt	tcattctgct	gaaacatttg
1921	aatcattact	ggcttgtctg	aaaatggatg	atgaaaaagt	agcagaagct	gcactacaaa
1981	ttttcaaaaa	cacaggaagc	aaaattgaag	aggattttcc	acacatcaga	tcagccttgc
2041	ttctgttttt	acatcacaaa	tctaaaaaag	gaccccccg	tcaagccaaa	tatgcccattc
2101	attgtatcca	tgcgatattt	tctagtaaag	agaccaggtt	tgacagata	tttagcctc
2161	tgcataagag	cctagatcca	agcaacctgg	aacatctcat	aacaccattg	gttactattg
2221	gtcatattgc	tctccttgca	cctgatcaat	ttgctgctcc	tttgaaatct	ttggtagcta
2281	ctttcattgt	gaaagatctt	ctcatgaatg	atcggcttcc	agggaaaaag	acaactaaac
2341	tttgggttcc	agatgaagaa	gtatctcctg	agacaatggg	caaaattcag	gctattaaaa
2401	tgatggttcg	atggctactt	ggaatgaaaa	ataatcacag	taaatcagga	acttctacct
2461	taagattgct	aacaacaata	ttgcatagtg	atggagactt	gacagaacag	gggaaaaatta
2521	gtaaaccaga	tatgtcacgt	ctgagacttg	ctgctgggag	tgctattgtg	aagctggcac
2581	aagaaccctg	ttaccatgaa	atcatcacat	tagaacaata	tcagctatgt	gcatttagcta
2641	tcaacgatga	atgctatcaa	gtaagacaag	tgtttgccca	gaaacttcac	aaaggccttt
2701	cccgtttacg	gcttccactt	gagtatatgg	caatctgtgc	cctttgtgca	aaagatcctg
2761	taaaggagag	aagagctcat	gctaggcaat	gtttggtgaa	aaatataaat	gtaaggcggg
2821	agtatctgaa	gcagcatgca	gctgttagtg	aaaaattatt	gtctcttcta	ccagagtatg
2881	ttgttccata	tacaattcac	cttttggcac	atgaccacga	ttatgtcaaa	gtacaggata
2941	ttgaacaact	taaagatggt	aaagaatgtc	tttggtttgt	tctggaaaata	ttaatggcta
3001	aaaatgaaaa	taacagtcac	gcttttatca	gaaagatggg	agaaaaatatt	aaacaaacaa
3061	aaagaccctg	aggaccagat	gatgacaaag	tgaatgaaaa	actgtacact	gtgtgtgatg
3121	ttgccatgaa	tatcatcatg	tcaaagagta	ctacatacag	tttggaatct	cctaaagacc
3181	cgggtactacc	agctcgtttc	ttcactcaac	ctgacaagaa	tttcagtaac	accaaaaatt
3241	atctgcctcc	tgaaatgaaa	tcattttttca	ctcctggaaa	acctaataaa	accaatgttc
3301	taggagctgt	taacaagcca	ctttcatcag	caggcaagca	atctcagacc	aaatcatcac
3361	gaatggaaac	tgtaagcaat	gcaagcagca	gctcaaatcc	aagctctcct	ggaagaataa
3421	aggggagggt	tgatagttct	gaaatggatc	acagtgaaaa	tgaagattac	acaatgtctt
3481	cacctttgcc	ggggaaaaaa	agtgacaaga	gagacgactc	tgatcttgta	aggtctgaat
3541	ttgagaagct	tagaggcagg	aaaaaaacgc	ccgtcacaga	acaggaggag	aaattaggta
3601	tggtagcact	gactaagttg	gtacaggaac	agaaaacctaa	aggcagtcag	cgaagtcgga
3661	aaagaggcca	tacggcttca	gaatctgatg	aacagcagtg	gcctgaggaa	aagaggctca
3721	aagaagatat	attagaaaat	gaagatgaac	agaatagtcc	gccaaaaaag	ggtaaaaagag
3781	gccgaccacc	aaaacctctt	ggtggaggta	cacaaaaaga	agagccaaca	atgaaaaactt
3841	ctaaaaaagg	aagcaaaaaa	aaatctggac	ctccagcacc	agaggaggag	gaagaagaag
3901	aaagacaaaag	tggaaatacg	gaacagaagt	ccaaaagcaa	acagcaccga	gtgtcaagga
3961	gagcacagca	gagagcagaa	tctcctgaat	ctagtgaat	tgaatccaca	cagtccacac
4021	cacagaaagg	acgaggaaga	ccatcaaaaa	cgccatcacc	atcacaacca	aaaaaaaaatg
4081	tccgtgttagg	acgctccaaa	caagcagcta	ctaaggaaaa	tgattcaagt	gaagaagtag
4141	atgtgtttca	gggtagctct	cctgtcgatg	atattccaca	ggaagaaaca	gaggaggagg
4201	aagtttctac	agtaaatgta	cggcggcgaa	gtgctaaaaag	ggaacggcga	tgaacaaatg
4261	taattaataa	ctttctctgt	gaaagctttg	gaaaaaatctt	tttttttttt	tttttttggt
4321	caagcttgag	gctgaataaa	gcctttgatg	cacaaaaatgg	gactgctgaa	gagtggacag

4381 ttggacctta ctttggtgac cccatacgtt tgtgggcaca tgcttttagcc atacacatgg
4441 taacattgac tatggagtct tgtgaaagtg taatgtgcga tggctatgta gacataaaga
4501 agaaacttgt aaatatcttt tttctttttt ttaatgtttc tgatttctga agtgcttgta
4561 tagcttttat ctgcggcttt aaactgacag taccgactg tttattggat ctattgattt
4621 gaaaagaatt tgtaggata gatcttaagc agtaatctgt cagtgtttgt atttgtattc
4681 tctgcaattt tactgtgaaa aaaaatttgt tttcaacaat tgggtgcatt ttcttgatgt
4741 cactatttgt tggagagtta aatggtctct tccctttgtg tatcttacct agtgtttact
4801 cctgggcacc cttaatcttc agagggtgcta aattgtctgc cattacacca gaaggatgcc
4861 tctgatagga ggacaaccat gcaaattgtg aaatagtcct gaagttcttg gattacttta
4921 cacctcagta ttgatttgtc ccagaatttt ctggcctttc atggcaatga aaattttaag
4981 aagaaagatt taaagtattt taatttttaa gagtgtgtta taaaataatg tactgaattc
5041 tttatcccat tttatcatcc tttcagtttt tattaatcta ctgtatcaat aaaattctgt
5101 aatttgaatg

=>

L1 ANSWER 1 OF 1 GENBANK.RTM. COPYRIGHT 2001

LOCUS (LOC): HSU50533 GenBank (R)
GenBank ACC. NO. (GBN): U50533
CAS REGISTRY NO. (RN): 180884-10-0
SEQUENCE LENGTH (SQL): 1852
MOLECULE TYPE (CI): mRNA; linear
DIVISION CODE (CI): Primates
DATE (DATE): 27 Nov 1996
DEFINITION (DEF): Human BRCA2 region, mRNA sequence CG008.
SOURCE: human.
ORGANISM (ORGN): Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Euteleostomi; Mammalia; Eutheria; Primates;
Catarrhini;
Hominidae; Homo
NUCLEIC ACID COUNT (NA): 609 a 348 c 379 g 505 t 11 others
REFERENCE: 1 (bases 1 to 1852)
AUTHOR (AU): Couch, F.J.; Rommens, J.M.; Neuhausen, S.L.; Belanger, C.;
Dumont, M.; Kenneth, A.; Bell, R.; Berry, S.; Bogden, R.;
Cannon-Albright, L.; Farid, L.; Frye, C.; Hattier, T.;
Janecki, T.; Jiang, P.; Kehrer, R.; Leblanc, J.-F.;
McArthur-Morrison, J.; McSweeney, D.; Miki, Y.; Peng, Y.;
Samson, C.; Schroeder, M.; Snyder, S.C.; Stringfellow, M.;
Stroup, C.; Swedlund, B.; Swensen, J.; Teng, D.;
Thakur, S.;
Tran, T.; Tranchant, M.; Welter-Feldhaus, J.;
Wong, A.K.C.;
Shizuya, H.; Labrie, F.; Skolnick, M.H.; Goldgar, D.E.;
Kamb, A.; Weber, B.L.; Tavtigian, S.V.; Simard, J.
TITLE (TI): Generation of an integrated transcription map of the
BRCA2 region on chromosome 13q12-q13
JOURNAL (SO): Genomics, 36 (1), 86-99 (1996)
OTHER SOURCE (OS): CA 125:187182
REFERENCE: 2 (bases 1 to 1852)
AUTHOR (AU): Simard, J.
TITLE (TI): Direct Submission
JOURNAL (SO): Submitted (04-MAR-1996) Jacques Simard, Laboratory of
Molecular Endocrinology, CHUL Research Center, 2705,
Boulevard Laurier, Quebec City, Quebec G1V 4G2, Canada

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..1852	/organism="Homo sapiens" /db-xref="taxon:9606" /chromosome="13" /map="13q12-q13" /note="CG008; DSEG numbers: D13S171 and D13S310"

SEQUENCE (SEQ):

```
1  ttaaagtgag gtcacatgac cctgaggaag ctattagaca tgatgttatt gtgtcaatag
61  ttacagctgc taaaaaggat attcttctgg tcaatgatca cttacttaat tttgtgagag
121 agagaacatt agacaaacga tggagagtag gcaaagaagc catgatggga cttgccccaa
181 tttataagaa atatgcttta cagtcagcag ctggaaaaga tgctgcaaaa cagatagcat
241 ggatcaaaga caaattgcta catatatatt atcaaaatag tattgatgat cgactacttg
301 ttgaacggat ctttgctcaa tacatggttc ctcaaatat agaaactaca gaacggatga
361 aatgcttata ttacttgtat gccacactgg atttaaagtc tgtgaaagca ttgaatgaaa
```

421	tgtggaaatg	tcaaaatctg	ctccgacatc	aagtaaagga	tttgcttgac	ttgattaagc
481	aacccaaaac	agatgccagt	gtcaaggcca	tattttcaaa	agtgatggtt	attacaagaa
541	atttacctga	tcctggtaag	gtcaggatt	tcatgaagaa	attcacacag	gtgttagaag
601	atgatgagaa	aataagaaag	cagttagaag	tacttgtag	tccaacatgc	tcctgcaagc
661	aggctgaagg	ttgtgtgctg	gaaataacta	agaagttggg	caaccccaaa	cagcctacaa
721	atcctttcct	ggaaatgatc	aagtttctct	tggagaggat	agcacctgtg	cacatagata
781	ccgaatctat	cagtgtctct	attaaacaag	tgaacaaatc	aatagatgga	acagcagatg
841	atgaagatga	gggtgttcca	actgatcaag	ccatcagagc	aggtcttgaa	ctgcttaagg
901	tactctcatt	tacacatccc	atctcatttc	attctgctga	aacatttgaa	tcattactgg
961	cttgtctgaa	aatggatgat	gaaaaagtag	cagaagctgc	actacaaatt	ttcaaaaaca
1021	caggaagcaa	aattgaagag	gattttccac	acatcagatc	agccttgctt	cctgttttac
1081	atcacaaatc	taaaaaagga	ccccccgct	caagccaaat	atgccattca	ttgtatccat
1141	gcgatatttt	ctagtaaaga	gacycagttt	gcacagatat	ttgagcctct	gcataagagc
1201	ctagatccaa	gcaacctgga	acatctcata	acaccattgg	ttactattgg	tcataattgct
1261	ctccttgcac	ctgatcaatt	tgctgtcctt	ttgaaatctt	tggtagctac	tttcattgyg
1321	aaagatcttc	tcatgaatga	tcggcttcca	gggaaaaaga	caactaaact	ttgggttcca
1381	gatgaagaag	tatctcctga	gacaatggtc	aaaattcagg	ctattaaaat	gatggttcga
1441	tggctacttg	gcatgtaamt	tamtcacagt	atatcaggat	cttcttcctt	aagsattgct
1501	aacaacaata	tcgcatcgtg	atggggattg	gmagamcagg	ggaaattagt	aaccagatat
1561	gtcacgtctg	agacttgctg	ctgggagtgc	tatttgaag	ctggcacaag	aaccctgtta
1621	ccatgaaatc	rtcacawtag	aacaatatca	gctatgtgcr	ttagctatca	acgatgaatg
1681	ctatcaagta	agacaagtgt	ttggcccaga	aacttcgcaa	aggcctttcc	cgtttgcggc
1741	ttccacttga	gaatatggca	atctgtgccc	tttgtgcaaa	agatcctgta	aaggagagaa
1801	gagctcatgc	taggcaatgt	ttggtgaaaa	atataaatgt	aaagcgggac	tg

=>